

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS



THE UNIVERSITY OF ALBERTA

STUDENT RESPONSES TOWARD AN ORAL
EXAMINATION AS A FUNCTION OF
FEEDBACK

BY



WINSTON F. YOUNG SOON

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

SPRING, 1971

Thesis
(2-1)
175
ii

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies for acceptance, a thesis entitled "Student Responses toward an Oral Examination as a Function of Feedback" submitted by Winston Frederick Young Soon in partial fulfillment of the requirements for the degree of Master of Education.

ABSTRACT

In this study, oral examinations were individually given to an undergraduate class of 43 students enrolled in an Educational Psychology course. For one group of subjects, verbal and nonverbal feedback (FB) was provided while feedback was withheld for the other group (NO-FB). On the basis of the results of a pilot study along with a review of the related literature, hypotheses were postulated in the following areas: learning, evaluation, atmosphere, anxiety, communication, and attitudes. Student responses were obtained through the use of a questionnaire as well as a semantic differential, both of which were administered after all the oral examinations were conducted. Each oral lasted approximately thirty minutes and was based on an essay each student had written. It was given by three examiners, the course instructor in addition to two graduate students, one of whom was the research worker.

It was found that Ss in the FB group tended to view the oral examination as more of a learning experience than the NO-FB subjects. However, support for this hypothesis was only evident from an analysis of the data from the semantic differential but not from the questionnaire. Subjects in the FB group also claimed that they experienced greater gain in terms of knowledge and better understanding of the implications of their essays than the Ss in the NO-FB group.

In comparison to those in the NO-FB group, Ss in the FB group rated the oral examination as a better evaluative tool than a written examination. They felt, too, that the oral provided them with a greater opportunity to discuss or defend their essays and gave them a better opportunity to explain more than they would normally have had in

a written examination. However the responses of the Ss in the FB group were significantly more positive than in the NO-FB group.

Compared to Ss in the FB group, Ss in the NO-FB group perceived the atmosphere during the oral as being unfriendly, impersonal, threatening, etc. Furthermore both groups felt that the atmosphere during the oral made a difference in their performance.

There was no significant difference between the two groups in their reported level of anxiety just before the oral examination, rated ex post facto. However Ss in the NO-FB group rated themselves as being more anxious during the oral than those in the FB group. They also felt that their anxiety adversely affected their performance. Analysis of the marks assigned by examiners and students revealed no significant difference between the two groups.

Both the FB and NO-FB groups rated themselves no differently in their ability to communicate verbally. In addition, both groups felt that their ability to communicate verbally assisted them equally well in the oral examination.

Finally, it was found that subjects in the FB group had more positive attitudinal responses toward the oral than those in the FB-NO group.

Correlation coefficients between the composite marks assigned by the three examiners exceeded 0.60. However, correlations were slightly lower when the marks on the criterion variables were considered. It was also found that correlations between examiners and students were very low, but positive.

ACKNOWLEDGEMENTS

I wish to express my sincerest appreciation to DR. METRO GULUTSAN for his most scholarly direction toward the completion of this thesis. His concern for my professional development together with his qualities as a human being have engendered an everlasting impression of what one man can be and is.

I would like to extend a warm thank you to PAUL PEEL whose enthusiastic participation and insightful comments truly reflect his worth as a person.

I would also like to thank DR. JOHN MCLEISH and PROFESSOR RYAN for their valuable suggestions which have made this thesis acceptable.

To Gloria McCormick and Karen Kauppila, thank you for your efficient typing.

And, finally, I wish to acknowledge my constant source of inspiration and love — my wife, Norma. This thesis is dedicated to her.

CONTENTS

CHAPTER	PAGE
I INTRODUCTION	1
Rationale and Purpose of the Study	1
II RELATED RESEARCH	5
On Testing and Examinations in General	5
On Oral Examinations	8
On Feedback	13
On Anxiety, Communication, and Performance	15
On Attitude and Attitude Measurement	18
III PILOT STUDY AND HYPOTHESES	22
Pilot Study	22
Hypotheses	22
IV PROCEDURES	28
The Sample	28
The Method	28
The Instruments	32
Analysis of Data	35
V ANALYSIS AND INTERPREATION OF RESULTS	38
Reliability of the Oral Examination	38
Preamble to Presentation of Results	40
Results	41
(a) Oral Examination as a learning experience	41
(b) Oral Examination as a means of evaluation	48
(c) Anxiety and the oral examination	50

CHAPTER	PAGE
(d) Atmosphere during the oral examination	55
(e) Communicative ability and the oral examination	60
(f) General attitude toward the oral examination	60
VI DISCUSSION	65
VII SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	73
Summary and Conclusions	73
Action Recommendations	75
Research Recommendations	76
SELECTED REFERENCES	77
APPENDICES	80

LIST OF TABLES

TABLE	PAGE
1. PEARSON PRODUCT MOMENT COEFFICIENTS OF CORRELATION OF GRADES ASSIGNED BY EXAMINERS AND STUDENTS ON THE ORAL EXAMINATION	39
2. SUMMARY OF COMPARISONS OF MEANS FOR FEEDBACK AND NO-FEEDBACK GROUPS ON THE QUESTIONNAIRE	42
3. COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK (NO-FB) GROUPS OF COMBINED MEANS FOR THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT 'EDUC. PSY. 471 ORAL AS A LEARNING EXPERIENCE.....	47
4. COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK GROUPS OF COMBINED MEANS FOR THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT 'EDUC. PSY. 471 ORAL AS A MEANS FOR EVALUATING STUDENTS'.	52
5. MEANS OF MARKS ASSIGNED BY EXAMINERS AND STUDENTS FOR FEEDBACK AND NO-FEEDBACK GROUPS	54
6. ANALYSIS OF VARIANCE OF MARKS CLASSIFIED BY MARKS ASSIGNED BY EXAMINERS AND STUDENTS FOR FEEDBACK AND NO-FEEDBACK GROUPS.....	56
7. COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK (NO-FB) GROUPS OF COMBINED MEANS FOR THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT 'ATMOSPHERE DURING EDUC. PSY. 471 ORAL'.....	59

TABLE	PAGE
8. COMPARISON OF FEEDBACK (FB) AND NO-FEEDBACK GROUPS OF COMBINED MEANS FOR THE EVALUATION DIMENSION OF THE SEMANTIC DIFFERENTIAL	63

LIST OF FIGURES

FIGURE		PAGE
1	PROFILE OF THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT, 'EDUC. PSY. 471 ORAL AS A LEARNING EXPERIENCE' FOR FEEDBACK AND NO-FEEDBACK GROUPS	46
2	PROFILE OF THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT, 'EDUC. PSY. 471 ORAL AS A MEANS FOR EVALUATING STUDENTS' FOR FEEDBACK AND NO-FEEDBACK GROUPS	51
3	PROFILE OF THE SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT 'ATMOSPHERE DURING THE ORAL' FOR FEEDBACK AND NO-FEEDBACK GROUPS	58
4	PROFILE OF SEMANTIC DIFFERENTIAL RESPONSES MADE BY FEEDBACK AND NO-FEEDBACK GROUPS TO THE SCALES BELONGING TO THE EVALUATION DIMENSION	62

CHAPTER I

INTRODUCTION

Oral examinations in North American education, for the most part, have been restricted to graduate degree programs and foreign language evaluation. The reluctance to use this type of evaluation at other levels reflects, perhaps, the difficulties encountered in trying to attain objectivity. However, it should be noted that any examination where the student is provided with an opportunity for self-expression, whether oral or written, is bound to be of the subjective type. Other reasons for not using the oral type of evaluation are: variations in scoring standards among examiners, fluctuations in marking by the same examiner, fluctuations in levels of difficulty from year to year, high enrollments, difficulties in scheduling, and questionable validity. According to Atiyeh (1969), some people rested their case against oral examinations on psychological grounds and charged them with failure to produce a suitable climate for healthy personality development: orals were the source of shattering strain and anxiety. Complaints also came from administrators and teachers. The former voiced dissatisfaction with the cumbersome machinery needed for arranging examinations; the latter complained that with crowded classes, examinations and scoring took too much time and interfered with teaching.

While it is true that any of the above reasons might be sufficient to discourage the use of oral examinations, it is the writer's contention that this kind of examination not only evaluates an additional dimension of the pupil's knowledge and understanding, but also

provides the student with an opportunity to learn from the kind of interaction that can only be achieved in an oral situation. It is also interesting to note that many of the arguments voiced against oral examinations are equally applicable to written examinations, including the objective type. Moreover, Atiyeh (1969) acknowledging the many faults of the examination system itself and looking at future prospects, observed that

In spite of all this, there is no indication, however, of a decline in the importance of exams in modern times. Quite the contrary, there is now what appears to be the beginning of an era when focus on examinations will be at a maximum [p.376].

He also feels that greater concern will be shown for an understanding of the technical nature of examinations and of the role which they should play in education. An inevitable extension of this concern ought to be more extensive research into the improvement of oral and practical examinations.

It might be argued that most of the learning in schools take place during the course of lectures, class discussions and seminars (but even then one cannot assume that this is so) and that the only purpose of oral examinations or for that matter, any examination, would be to evaluate the learning that has already taken place. In contesting such an argument, one might say that the very fact of high enrollments (a reason cited against the use of oral examinations) limits class discussion and in fact much of the learning that occurs, comes from the student's own readings and investigations. Should this be

the case, the oral examination, properly conducted, can be a learning experience for the student in that it provides feedback and aids in clarifying previously vague or misunderstood concepts that may have developed during the course of study.

Because most of the instruction taking place in schools involves the use of language and often requires some interaction between teacher and student, it is seemingly justifiable to use oral examinations which involve both language and interaction, as part of the total process of evaluation. Besides, there is need to improve the oral communication skills of students; it becomes a matter of concern when the student's first oral presentation or examination comes only in the final year of a university program or not until graduate school. In addition, oral interaction makes it possible for the student to defend or explain further what he is really trying to communicate, and in so doing provides a more adequate means of assessing what he is capable of doing. And finally, the student, forced to verbalize in an oral situation, soon learns that communication requires organization of thought as well as clarity of presentation, attributes which ought to be part of the goals of education.

At the undergraduate level, students have frequently complained about the impersonal relationships between instructors and themselves. Partly as a result of this, they have come to view their education at university as meaningful in so far as they have gained technical competence and improvement in their job-seeking potential, but meaningless in terms of interpersonal growth as well as growth in affective and

cognitive ways. While there are many factors that may account for this, nevertheless, educators must strive to create a more humanizing environment as one step toward bridging the gap between the different hierarchical levels created within our institutions. The obvious implication is for a movement toward closer interaction among all the participants in institutions of learning. However, it is important to recognize that mere interaction is not enough; it is necessary to understand the nature of the interactive process as well as the conditions which foster the development of positive attitudes.

This study has arisen because the author felt that the oral examination has been neglected in favor of more expedient but not necessarily more sufficient ways of evaluating students. Since one of the primary concerns of educators should be the continuing improvement of the educational environment, any step toward improvement must, of necessity, deal with the means by which educators can best determine whether the objectives of education are being met. Therefore, this study has focussed upon the oral examination in an attempt to study some of the variables that impinge upon the successful evaluation of students; and specifically, how the students react to the kind of feedback given during the oral examination.

CHAPTER II

RELATED RESEARCH

On Testing and Examinations in General

Chambers (1964) made the following comment regarding testing and its relationship to educational objectives:

Although it would be generally agreed that tests definitely have a direct influence upon teaching and learning activities, little has been done to insure that these instruments continue to lend support to the objectives of education rather than to direct them.... The fact that the test may not be measuring the same learning process as those that are outlined in the school's objectives, is frequently overlooked [p. 246].

It is the writer's opinion that if the majority of items on examinations encourage the student to utilize his intellectual abilities, freely and flexibly, the examination will probably have a constructive influence on the learning process. However if the emphasis is placed upon rote learning, the examination may be a deterrent to learning experiences which could otherwise have real meaning and purpose for the student. A properly planned oral examination, that is, one that relates well to the objectives of education, can have just such a constructive influence. Chambers further stated:

However important the acquisition and retention of facts may be, it is doubtful that administrators and teachers would ever agree that their efforts in working with students are primarily directed toward these objectives [p.246].

Yet, in reviewing evaluative instruments - whether they be teacher-made, standardized, objective, or essay type, or whether this review

takes place on the elementary or graduate school level - the writer comes away with the impression that all that is required on examinations are the facts. It is no surprise, therefore, to find that students have a rather narrow viewpoint regarding the ideology of the educational system.

Educational objectives are often carefully formulated around terminology indicating the desirability of providing an environment conducive to the development of student abilities in such areas as knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom, 1956). However these well-worded objectives are often unrelated to evaluative instruments that purport to measure them. It is Chambers' thesis that these objectives have been delineated, within broad limits, by educators at all levels; that it is now time to concentrate efforts toward the development of tools and procedures which will, in fact, actually evaluate the degree to which these objectives are realized.

Oppenheim, Jahoda and James (1967) have questioned the assumption that examinations can include some so-called imponderables such as 'quality of mind', 'independent critical thinking' and 'breadth' in their assessment. Behind the differences in content of each examination there appears to be a common double purpose: to assess the student's knowledge and quality of mind. Unfortunately both these terms are liable to vastly different interpretations. Unless examiners specify clearly which interpretations of knowledge and quality of mind they wish to emphasize, it will be difficult to investigate whether examinations actually do test knowledge and quality of mind

nor will it be possible to increase the reliability of examination grading. This is even more relevant for oral examinations, since the questions often asked cannot always be prepared in advance, but may depend on the extent and kinds of interaction between examiners and examinees.

Apart from the difficulty of defining terms and hence of assessment, the question arises whether 'quality of mind' is at all influenced in a purposeful way by formal education. According to Oppenheim et al (1967) it is just possible that, could it be assessed, it would be much the same in the student when he is admitted and when he sits his final examination. Many teachers assume that to change a quality of mind in the process of education is impossible or not part of their task, and that changes which do occur are the result of maturation or that they occur randomly. This raises the vital question: is education really possible? For others the opposite position is taken (and seems implied by their use of the phrases, 'learning to learn' or 'learning to think') and the improvement of the quality of mind is made a goal of education. The point is relevant to examinations to the extent that examiners insist that quality of mind is a matter to be assessed. Moreover, the evaluative instruments that are presently available are based on the assumption that the answers given in response to questions asked reflect to some degree, or are isomorphic with, quality of mind. Given this assumption it remains to decide which instruments best perform the purpose of evaluation.

Ager and Weltman (1967) stated "There is, in fact, an increas-

ing awareness that no single examination is completely satisfactory in terms of both reliability and validity [p. 272]." On the one hand there is the objective type test which can be applied only to a limited range of intellectual skills; on the other, the essay and oral types which bear a more valid relationship to the aims of education, but which cannot be marked with complete objectivity. The point to be made is that a variety of techniques should be used in examinations, such techniques being chosen according to the functions that the examination best serves. The leading French expert on docimologie, the science of examining, Pieron (1963) sums up the position as far as reliability of traditional examinations is concerned:

All the experimental data has shown that for a particular performance expressed in terms of an exam script, assessment by different examiners produces marks with considerable variability, such that in the determination of these marks, the part played by the examiner can be greater than that of the performance of the examinee (translation in R. Cox, 1969, p. 75).

On Oral Examinations

In most countries of Europe, Asia and Africa, there is a strong tradition of "subjective" examining (essay type or oral). The form of the examination may, however, differ from one country to another. For instance, in England, India, Egypt, and Jordan, the trend is for written examinations. In East and West Germany and East European countries, the emphasis is on orals. In the Soviet Union, apart from a written examination in literature, language and mathematics, all subjects covered in the final examination are examined orally. In other

Middle Eastern countries, both written and oral examinations are used, with slightly greater emphasis on the written. It may seem surprising in view of the criticisms about subjectivity, reliability, etc. that up to the present, oral examinations have not fallen into disrepute. In fact they are still widely used, not only in languages, but also in a variety of subjects.

Perhaps the most widespread use of oral examinations exists in Denmark (Graham, 1963) where they are first used at the end of the sixth grade. These examinations are administered individually in designated fields of study which vary from year to year and include some subjects in which pupils have already had long written tests. School authorities make the decision as to the subjects in which a class will be examined orally. All examinations are conducted by the classroom teacher and a censor who is a visiting teacher from the same or another school.

Danish teachers regard the examinations as an important part of the educational process. They believe that the oral serves as a powerful stimulus for independent study and also encourages students to express themselves effectively in oral communications. According to Graham (1963) the relatively lengthy session adds a dimension to evaluation that is more penetrating than spot questioning in classroom situations and the usual written tests. It also provides for interaction between teacher and pupil and helps the teacher determine what a pupil has misconstrued or has failed to grasp.

Brereton (1969) reported that in the German Democratic Republic candidates taking oral examinations are allowed to see the questions

they are to answer half an hour in advance. They sit with no books under supervision in another room, with the opportunity to assemble their thoughts on paper before they face their examiners.

Morrissett (1958) used oral examinations in several undergraduate courses and reported the following generalizations: a large majority of the students preferred oral to written examinations; most felt that oral compared to written examinations provided them with a greater incentive to study, that they covered at least as much ground, and that they gave the student a better chance to show his knowledge of the course. He saw no reason for not accepting students' claims that they learned more from oral than from written examinations. A group oral examination was used where four students were examined together for a period of one hour. There were three such examinations in each semester. An interesting sidelight of the group examinations is the diplomacy it brings out: although most students were striving to do well, they did not wish to discredit the efforts of a fellow student.

Barnes and Pressey (1929) correlated the grades assigned by different pairs of examiners (there were three) and found that while correlations on individual traits by which the students were examined were low, the correlations for the composite ratings were relatively high. They also compared the grades obtained on the oral examination with those made on an objective type examination and found that the validity coefficients were relatively low. They felt that the ratings on the oral examination and the scores on the objective test represent different functions but were unable to define what these functions were.

Trimble (1934) utilized the oral examination as a final examination. Students were asked to select what they individually considered to be a major aspect or subject of their course and were told to prepare for thorough examination on it. Each student was examined for only twelve minutes and assessed on vocabulary, knowledge and general ability. He found correlations of 0.333, 0.404, and 0.225 between the grades assigned by three examiners.

Cox (1969) cited the classic study done by Hartog and Rhodes (1935) in which the marks given by two very experienced examining boards were compared. The candidates were fairly homogeneous, therefore, a correlation coefficient between the two sets of marks was equally low, 0.41.

Carter (1962) analyzed data from a sample of 250 candidates examined orally in a field of medical specialization and found evidence of high reliability of a set of oral examinations, and evidence of only moderate agreement between measures provided by written and oral examinations. He suggested that in view of the low correlation with a written examination, the oral examination might have a unique function. That is, it might result in an assessment of aspects of competence not adequately covered in written examinations. However, like Barnes and Pressey, he failed to specify what these aspects were.

Narciso (1964) conducted oral examinations which the students themselves administered to each other. The class was divided into groups of four students who took turns in giving and answering questions. He found that this type of oral examination yielded positive attitudinal results. Furthermore the procedure yielded quantitative

data which correlated well with a standard objective-type testing method.

Evidence obtained by Pieron (1963) and cited in Cox (1969) from the 1955 French Baccalaureat showed that the variability from one examiner to another was, on the whole, greater in oral examinations than in written. Even when the same oral examination is judged from a recording by a number of examiners, the reliability is still greater with a written examination. However he concluded that the relative value of written and oral examinations was still an open question.

Cox (1967) felt that the difficulty with regard to oral examinations was that most examiners took a rather optimistic view of their own capabilities. On the whole the danger from 'halo' effects was high and the intrusion of irrelevant personality factors was difficult to eliminate. Moreover, judgments were often made not on knowledge or acquired abilities but rather on alertness, intelligence and general outlook.

From a review of the literature on oral examinations then, the following conclusions about previous research were derived: there appear to be differing degrees of reliability and validity found by different researchers. Some of these differences may stem from the varying examining techniques used as well as from the comparisons made between scores on oral and written examinations. The assumption here is that they are evaluating the same things. Almost all research workers investigating attitudes found fairly positive student reactions to oral examinations. However there is a marked dearth of research

concerning comparisons between different types of oral examinations, the nature of the interaction between examiners and examinees and the resulting influence on the attitude of students toward this kind of examination.

On Feedback

The concept of feedback has been proposed as a variable having potentially the highest significance in the production of learning. McKeachie (1962) repeatedly uses the concept, both in setting out the theoretical differences between the methods compared in much previous research (lecture-discussion, autocratic-democratic, teacher oriented-student oriented), and in helping to explain the results obtained. His and other discussions tend to centre around feedback as utilized by the individual student.

Solomon and Rosenberg (1964) defined feedback as the

...mechanism by which a machine, device or organism receives information about the nature and/or effects of its behavior. It is usually conceived as part of a homeostatic process; the feedback is utilized to determine whether a course of action is the correct one, or whether some alteration is necessary to make it correct (p. 197).

According to Watzlawick, Beavin, and Jackson (1967) interpersonal systems e.g. families, stranger groups, participants in oral examinations, may be viewed as feedback loops since the behavior of each person affects and is affected by the behavior of every other person. Input into such a system may be amplified to produce change or may be counteracted to maintain stability, depending on whether the feedback mechanisms are positive or negative.

Elements of feedback are contained implicitly in certain classical psychological concepts. Reinforcement, for example can be thought of as providing feedback to the individual about some of the effects of his behavior and as fulfilling an orienting function. However feedback is not limited to the presentation of reinforcement. Often it provides additional information which identifies precisely, the direction and extent of the error. This is not different from other conceptions of feedback. For instance, both reinforcement and specific information are involved in the feedback built into programmed instruction.

In an oral examination an individual can receive feedback from several sources: from his own ongoing behavior, from his own judgment of its effects, and from the nonverbal cues and comments given by the examiners during the course of behavior and following it. According to Solomon and Rosenberg (1964) the examiner's feedback may serve to reinforce a narrow or broad aspect of the total behavioral field. It may refer to the correctness of the information, to the general approach the student is making to the question (e.g. succinct v.s. discursive), or simply to the fact that he is responding. The examiners may also give differential reinforcement to spontaneous as opposed to solicited comments, to directly relevant material as opposed to peripherally relevant material, and to facts rather than opinions. However, even if an examiner provides a great deal of feedback, the student may still be in doubt as to his relative knowledge, ability and performance especially if the feedback consistently contains little information.

The effectiveness of feedback also depends on the student's perception as to the examiner's competence (does he believe that the examiner is adequately prepared), legitimacy (is he commenting on an appropriate or relevant topic), frankness (does the examiner never make a negative comment), and the manner in which the feedback is given (is the examiner accepting). Solomon and Rosenberg further state that comments made by the examiner regarding areas in which he is seen as an expert are likely to have great salience and to produce relatively strong effects. But feedback which the student for one reason or another, believes to be inappropriate or threatening would likely have very little effect. In either case, attitudes and opinions will inevitably be formed by the student and the extent to which he adopts a positive or negative response to the oral situation depends on many factors, some of which are mentioned above, and will most certainly influence the process of learning.

In a study carried out by Leavitt and Mueller (1951) it was found that feedback, freely given, improved accuracy in interpersonal communication; it also developed confidence and amity. The lack of feedback brought about the opposite effect.

On Anxiety, Communication, and Performance

When interacting in an interpersonal situation, one may respond to any or all of a number of cues. Most often, individuals respond directly to the content of another's message. That is, individuals are conscious of the words communicated and believe they are dealing with other persons at the level of linguistic messages. Shapiro (1968) has shown, however, that individuals are responsive to other than

linguistic cues; one's personality, emotional state, and other variables (e.g. atmosphere) are particularly relevant.

Eisenson, Aver, and Irwin (1963) presented two views of the relationship between emotions and behavior. First, emotions, e.g. anxiety and fear, are seen as disruptive and disorganizing forces, which impair rather than help the individual in his effort at adjustment. Exponents of this viewpoint argue that emotional responses are characterized by predominantly random, excessive, and largely useless motor as well as verbal behavior. An alternate view of emotional behavior emphasizes the reorganization rather than the disorganization of response patterns:

An individual tends to meet "new" situations in terms of his background of experience. If he finds that his experiences are inadequate, or inappropriate, he attempts to meet the new situation with a modified response pattern, a new approach. . . If he finally perceives that his new approaches are without avail, and yet feels driven to do something either because he cannot or will not extricate himself from the situation, then he may give way to disorganized behavior (p. 71).

Eisenson et al. (1963) further discussed emotional behavior as evidenced by fear in known or unknown situations. They claim that the common and essential component in fear-provoking situations, regardless of whether they are familiar or unfamiliar, is the assumption or the realization by the individual that he is inadequate. When an individual thinks that he lacks the power or the ability to deal effectively with a situation he is about to confront or one in which he is immediately involved, fear is evoked. It seems obvious that the

examination is a potentially fear-inducing situation and the results of such fear can have adverse affects on communicative ability. Yet, an optimistic view of the possibility of reducing fear and its effects is held by Eisenson and his colleagues as seen in the following comment:

It follows from this explanation of the dynamics of fear that training and experience prepare the individual to meet potentially fear-evoking situations with patterns of responses that eliminate or reduce the likelihood of fear, or at least the intensity of a fear reaction that is disruptive . . . (p. 73).

Although many studies of the relationship between anxiety and performance have been concerned with learning tasks (Mandler and Sarason, 1952; Raymond, 1953) there has been some interest in the effects of anxiety on speech (Benton, Hartman and Sarason, 1955). Since higher levels of stress and anxiety tend to foster disorganization, the language of the individual under stress would become disrupted and characterized by fragmentary speech and other signs of disorganization (Hassol, Cameron and Margaret, 1952). Thus the efficiency of communication is lowered (Gynther, 1957). Also serving to lower communicative efficiency would be verbal responses elicited by anxiety itself as well as incorrect responses and unanswered questions.

These findings are particularly relevant to the oral examination situation because of the relatively high anxiety generated by this procedure for some students and the individual differences that are obviously apparent. Hence research utilizing an oral examination

must control or match for these variables, especially if performance is being used as a dependent variable. Findings about the effects of anxiety on speech behavior take on added significance in relation to the fact that verbal communication is the basic medium through which responses are made by the student in an oral examination.

On Attitude and Attitude Measurement

Various definitions of attitude have been proposed, among which are the following:

An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related (Allport, 1935, p. 810).

An attitude is. . . "an enduring organization of motivational, emotional, perceptual, and cognitive processes with respect to some aspect of the individual's world" (Krech and Crutchfield, 1948, p. 152).

An attitude is a "relatively stable affective response to an object" (Rosenberg, 1956, p. 367).

An attitude is a "tendency or disposition to evaluate an object or the symbol of that object in a certain way" (Katz and Stotland, 1959, p. 428).

McGuire, in the 1968 Handbook of Social Psychology, uses Allport's (1935) definition to indicate five dimensions of disagreement among definitions of attitude. First, there is disagreement about the psychological locus of attitudes. Allport stated that attitudes were "mental and neural states". Krech and Crutchfield (1948) defined it as motivational, emotional, perceptual, and cognitive. A second dimension

of disagreement concerns whether attitude should be defined as a response or as a readiness to respond. A third issue concerns the degree to which attitudes are organized. To what extent is a single attitude made up of separate components? And to what extent are these components organized in the same way for attitudes toward different objects or situations? A fourth area of disagreement among definitions concerns the extent to which attitudes are learned through previous experience. A fifth issue concerns the extent to which attitudes play a directive-knowledge or a dynamic-motivational function.

There is some consensus that attitudes are learned and implicit - they are predispositions to respond but are unique in that they predispose toward an evaluative response. Hence attitudes are tendencies of approach or avoidance. It follows then that attitudes can be ascribed to some basic polar continuum with a neutral or zero reference point, implying that they have both direction and intensity, thus providing a basis for quantification. Consistent with this approach is Osgood et al's (1957) theoretical discussion of attitude and attitude measurement which they sum up in the following way:

This characterization of attitude as a learned, implicit process which is potentially bipolar, varies in its intensity, and mediates evaluative behavior, suggests that attitude is part. . . of the internal mediational activity that operates between most stimulus and response patterns (p.190).

If attitude is indeed some portion of the internal mediational activity, it is, by inference from our theoretical model, part of the semantic structure of an individual, and may be correspondingly indexed. The factor analysis of meaning may then provide a basis for extracting this attitudinal component of meaning (p. 190).

Within Osgood's framework, if an attitude is a predisposition to an evaluative response, it seems reasonable to identify it with the evaluative dimension of the semantic space which has been described as:

. . .a region of some unknown dimensionality and Euclidian in character. Each semantic scale, defined by a pair of polar (opposite-in-meaning) adjectives, is assumed to represent a straight line function that passes through the origin of this space, and a sample of such scales, then, represents a multidimensional space. The larger or more representative the sample, the better defined is the space as a whole (p. 25).

Since meaning has been defined as a point in the multidimensional semantic space, so too can attitude be defined as the projection of this point onto the evaluative dimension of that space. And since every point in semantic space has an evaluative dimension, every concept must, therefore, have an attitudinal component as part of its total meaning. To index attitude, then, it is necessary to procure a set of scales loading high on the evaluative factor but with negligible loadings on other factors. This can be done through factor analysis.

Although McGuire (1968) contends that there is relatively little agreement on the distinctions that have been made to differentiate attitude from other psychological concepts like beliefs and opinions, Eisenson, et al (1963) have provided a fairly acceptable approach to this problem. They stated that:

In contrast to the generalized nature of an attitude, a belief is the acceptance of a specific proposition.

An individual may have a general attitude of religious tolerance, but be quite categorical about the question, "Do you believe in God?" The establishment of a belief involves some thinking on the part of an individual; this is not necessarily true about an attitude. Individuals verbalize their beliefs and prefer to be able to state supporting reasons (p. 233).

They further commented that while attitudes and beliefs inevitably tend to interact, sometimes making them difficult to distinguish from each other, opinions are clearly responses derived from a combination of attitudes and beliefs. In practice, an opinion is an overtly verbalized attitude or belief, sometimes both. Like a belief, it is a statement of a proposition; like an attitude, it is favorable or unfavorable.

This discussion of the distinctions between the concepts, attitude, beliefs, and opinions, is relevant to this study concerning oral examinations to the extent that all three concepts constitute what has been called "responses", and have sometimes been used interchangeably. It is, however, necessary to distinguish between them.

CHAPTER III

PILOT STUDY AND HYPOTHESES

Pilot Study

In the fall of 1969, a pilot study was conducted by the writer in which 37 students were individually examined orally as one part of their course evaluation. There were 3 examiners: the course instructor and 2 graduate students, one of whom was the writer. Each oral examination lasted twenty minutes and consisted of questions derived from an essay the student had written. Examination conditions were similar to the feedback condition utilized in the present study (see p. 29). On the basis of the responses made to a questionnaire (see Appendix A) prepared by the writer it was found that a majority of the students:

- (a) favored oral examinations as one means of evaluation.
- (b) perceived a positive relationship between the atmosphere during the oral examination and performance.
- (c) would like to see the oral examination lengthened to at least a half hour.
- (d) felt that the questions were stimulating and fair.
- (e) viewed the oral examination as a learning experience.

About half the students felt that their anxiety adversely affected their performance.

Hypotheses

An examination of the findings obtained on the pilot study along

with a review of the related literature led to a deeper appreciation of the variables that impinged upon the oral examination itself and of the variables that influenced student's responses to the oral examination as a learning experience and an evaluative instrument.

There were two significant modifications in the design of the actual study - the addition of a comparison group characterized by a no-feedback condition (see p.29), and the use of two response measures, a questionnaire and a semantic differential. Both of these instruments gave quantifiable results.

Hypotheses were developed in the following major areas of the study: learning, evaluation, anxiety, atmosphere, communication and attitude.

(a) Oral examination as a learning experience

Examinations serve to evaluate a student's progress. While the oral examination can also achieve this function, it has the added advantage of providing a setting where interaction between examiners and examinees is possible. This unique feature of the oral examination, in combination with the utilization of feedback, could provide additional cues to the student which would inform him how well he was functioning and the areas in which he was deficient. The use of a group in which feedback is withheld emphasizes that the oral examination can serve as an examination only and need not be a learning experience for the student.

In comparison to Ss in the no-feedback (NO-FB) group, Ss in the feedback (FB) group would:

Hypothesis 1 = view the oral examination as more of a learning experience.

Hypothesis 2 = experience greater gain in terms of knowledge as a result of having the oral examination.

Hypothesis 3 = feel that they had gained more in terms of better understanding of the implications of their essays.

(b) Oral examination as a means of evaluation

The relative merits of written compared to oral examinations are still undecided. It is not the purpose of this section of the study to defend or refute claims for the alleged advantages of written examinations, but to determine how an oral examination can provide an opportunity for the student to function in ways which he could not do otherwise. It is the writer's contention that the no-feedback condition utilized in this study is not, in effect, very different from a written examination and would be so perceived by students. It was predicted that Ss in the feedback group would perceive the oral examination to be a better evaluative tool than a written examination.

In comparison to Ss in the NO-FB group, Ss in the FB group would:

Hypothesis 4 = rate the oral examination as a better evaluative tool than a written examination.

Hypothesis 5 = feel that the oral examination provided them with a greater opportunity to discuss or defend their essays.

Hypothesis 6 = feel that the oral examination gave them an opportunity to explain more than they would have normally communicated in a written examination.

(c) Anxiety and the oral examination

That the oral examination is an anxiety-arousing situation cannot be denied. It is expected that anxiety would adversely affect the performance of some students; but could serve to heighten the ability of others to focus upon the subject matter under consideration. Since subjects were randomly assigned and since they were unaware there were different "treatments," it was predicted that there would no significant difference in the reported anxiety level of the two groups just before the oral examination. However it was also predicted that there would be differential student reactions because of the different treatment conditions, with the Ss in the no-feedback group reporting significantly greater adverse effects on their performance as a result of anxiety.

Compared to Ss in the FB group, Ss in the NO-FB group would:

Hypothesis 7 = show no difference in their anxiety level
just before the oral examination.

Hypothesis 8 = rate themselves as being more anxious
during the oral examination.

Hypothesis 9 = feel that their anxiety adversely
affected their performance.

(d) Atmosphere during the oral examination

Although the notion of "atmosphere" seems to be an intangible concept describing an aspect of certain nonverbal kinds of interactions, nevertheless the fact that it is difficult to assess, does not preclude its influence on an individual's perception of a particular situation. It was predicted that a situation in which feedback is withheld would appear to be unfriendly, cold, threatening, unpleasant,

etc. The opposite would be true for a situation in which feedback is given. A further prediction was concerned with the effect such a perception would have on a person's performance.

Compared to Ss in the FB group, Ss in the NO-FB group would:

Hypothesis 10 = perceive the atmosphere during the oral examination to be unfriendly, cold, threatening, unpleasant, etc.

Hypothesis 11 = feel that the atmosphere during the oral examination made a difference in their performance.

(e) Communicative ability and the oral examination

The ability to communicate verbally is a function of previous learning as well as the conditions under which it is elicited or evoked. Since Ss were randomly assigned to their respective groups, there should be no difference in their ability to communicate verbally. However it was predicted that feedback (as defined in this study) would create the conditions that would facilitate verbal communication, whereas no feedback would inhibit it.

In comparison to Ss in the NO-FB group, Ss in the FB group would:

Hypothesis 12 = rate themselves no differently in their ability to communicate verbally.

Hypothesis 13 = feel that their ability to communicate verbally assisted them more in the oral examination.

(f) General attitude toward the oral examination

The acquisition of an attitude toward a person, object, or

situation is dependent upon many factors, including the environment in which the individual is working and how well he performs in it. The person who perceives a situation as being hostile, threatening, etc. conceivably acquires unfavorable reactions even if he performs adequately in such a situation. However, if the situation is a friendly, helping, pleasant one, the person can still adopt a favorable stance to it even if he knows that his performance was inadequate. It becomes relevant to the extent that the person approaches or avoids such situations in the future and to the extent it affects subsequent learning.

Hypothesis 14 = In comparison to Ss in the NO-FB group, Ss in the FB group would have more positive attitudinal responses toward the oral examination.

CHAPTER IV

PROCEDURES

The Sample

The sample consisted of an undergraduate class (43 students) enrolled in a third year course in Educational Psychology. All students had had at least one previous course in Psychology or Educational Psychology. The majority were either in their second, third or fourth year of a basic degree program in Education. A few belonged to the Faculty of Nursing and a small number were also enrolled in the post-degree course in Education.

There were 21 students in the group receiving feedback and 22 in the group in which feedback was withheld. A stratified random sampling procedure was utilized so that the number of males and females in each group was approximately the same. There were 13 females in each of the two groups, but only 8 males in the feedback (FB) group and 9 in the no-feedback (NO-FB) group.

The Method

At the beginning of the term the students were informed that an oral examination would be based mainly on an essay which was to be handed in prior to the oral examination. Selection of a topic for the essay was to be made from a list of several areas related to the course. In the second week of the course each student was given a short interview (10-15 minutes) in which his choice of a topic was discussed with one of the three examiners. This also served as an introduction to the setting for the oral interview situation for

all interviews took place in the office in which subsequent oral examinations were held. Students were also informed that the oral examination constituted only a part (20%) of the total course evaluation, which also included a mid-term examination, the essay, and a final written examination.

A committee of three conducted the oral examinations. This committee was the same one that had given the oral examination with a different class, the previous term, for the pilot study. The examiners were briefed about their roles in each of the FB and NO-FB conditions. A definition of these roles is presented as follows:

Role of the Examiners in the NO-FB Condition

In this condition, the examiners:

- (a) asked questions but did not provide any verbal feedback as to the correctness of the answers,
- (b) attempted to refrain from providing non-verbal feedback,
- (c) avoided answering any questions that the student asked,
- (d) attempted to be objective in their evaluation without trying to increase tension or raise the students' anxiety,
- (e) attempted to limit the interaction to questions and answers,
- (f) did not follow up a student's response. It was up to the student to provide an adequate answer.

Role of the Examiners in the FB Condition

In this condition, the examiners:

- (a) asked questions and provided feedback,

- (b) answered all questions that the student raised,
- (c) encouraged maximum interaction through verbal and non-verbal means,
- (d) encouraged the student to discuss the implications of his essay for his own teaching,
- (e) attempted to develop rapport with the student by being friendly and supportive.

In an effort to keep the examiners constantly aware of the different roles each treatment condition demanded, a self-rating sheet (see Appendix B) was given to each examiner before the oral examination and was filled out at the end of every oral. In it, the examiners were asked to rate how successful they were in maintaining the main requirements of each condition.

Students were told that the oral examinations were to be recorded on tape so that if there were any questions as to the grades assigned, they would be given an opportunity to hear the recording and an explanation as to how the grades were designated. No requests were made for this service. They were also informed that only the examiners would use the recordings. Anonymity was assured.

A list of the names of the students in their respective treatment conditions was prepared beforehand. Order of appearance depended on the time each student had indicated was favorable for him. A period of about thirty minutes was spent with each student. After every examination, the examiners graded the student without consulting one another on the following items:

- (i) Ability to report facts clearly.
- (ii) Understanding of concepts in essay.

- (iii) Ability to present summary of essay.
- (iv) Ability to argue (support or refute) topics reported in essay.
- (v) Ability to discuss relevance of topic for education (both theoretical and practical implications).

Each of the above variables was graded on a five-point scale. The student's composite score was an average of the scores assigned by the examiners. Students were also given an opportunity to grade themselves using rating sheets similar to the ones the examiners used. They were told that this information was need to compare their own ratings with that of the examiners, but that it would have no influence on the grades they received.

The oral examinations were conducted over a period of one week. A questionnaire and the semantic differential were administered to the entire group at one class sitting about two days after the last oral examination was completed. Prior to its administration, students were informed that the data obtained on these two instruments were needed to help find out more about students' reactions to the oral examinations. They were asked to read the instructions and were given as much time as was necessary to respond. They were requested to sign their names on both instruments and a promise was made that no one would look at their responses until their final grades for the course were submitted to the Registrar's Office at the university. They were also told that a summary of the findings would be mailed to each student when the study was completed.

The Instruments

Data on student reactions to the oral examinations were obtained through the use of two instruments:

- (a) A questionnaire (see Appendix C) which was prepared by the writer. It was felt that much of the information required to test the hypotheses could best be obtained by having the students respond to direct questions on a self-rating instrument whose scales ranged from 0 to 4. There was no reason for the writer to question the assumption that the responses made by the students on the questionnaire were reliable. Many of the questions originated from the questionnaire used in the pilot study.
- (b) The semantic differential (SD). Students were also asked to respond to a SD (see Appendix D) which included the following parts:
 - (i) Oral examination (with reference to the one they had all received),
 - (ii) The Educ. Psy. 471 Oral as a learning experience,
 - (iii) The Educ. Psy. 471 Oral as a means of evaluating students,
 - (iv) Atmosphere during the oral examination.

The researcher felt that the SD provided information which supplemented the data from the questionnaire. A number of the hypotheses were tested by analysis of data from both instruments. The SD also serves a second important purpose in that it was an indirect test of the reliability of students' responses; if the responses made on the questionnaire were reliable then they should agree with the responses elicited on the SD.

The usual SD consists of a number of scales or adjective pairs, chosen for a particular research problem, and the concepts to be rated. The scales are usually seven-point rating scales, the underlying nature of which has been determined empirically. That is,

each scale measures one, sometimes two of the basic dimensions or factors that Osgood and his colleagues found when they factor-analyzed the responses made on the SD. They discovered that adjective pairs like good-bad, bitter-sweet, clean-dirty, personal-impersonal and so on, fell into 3 main clusters. The most important cluster seemed to consist of adjectives that are EVALUATIVE, such as good-bad, pleasant-unpleasant, like-dislike, and fair-unfair. A second cluster had adjectives that seemed to share strength or POTENCY ideas, e.g., strong-weak, rugged-delicate, and tall-short. A third important factor was called ACTIVITY because its adjectives seemed to express motion and action, e.g., fast-slow, hot-cold, active-passive, and exciting-depressing.

The key question is how to define and identify clusters of adjectives. If correlation coefficients between adjective-pairs are used then the answer is factor analysis or some other simple form of cluster analysis. An additional procedure would involve oblique rotations of the axes of the semantic space; this would have the added advantage of selecting the most appropriate dimensions or factors so that they could more easily be identified. In this study, a number of adjective pairs, such as warm-cool, active-passive, friendly-unfriendly, harmonious-discordant, supportive-threatening, etc. whose correlation coefficients were factor analyzed, belonged to the same factor which was then labelled ATMOSPHERE.

The first step in constructing or selecting a SD for research use is to choose the concepts or other stimuli, e.g., religion, school,

examination, etc., one is going to rate with the adjective-pairs. One has to select a number of concepts that are conceived to be relevant to the research problem. A sample of concepts must be judiciously chosen to represent some parts of the semantic space. If the concepts used are not loaded with potential for identifying different reactions in people holding different attitudes, they are useless for measurement purposes. In short the concepts must be capable of eliciting varied responses and thus large variance.

The second step in the construction of the instrument is to select appropriate scales or adjective pairs. Two main criteria determine the selection: factor representativeness and relevance to the concepts used. In most cases, the research worker will want to use scales representative of the three main factors - evaluation, potency and activity. It is possible however that he may wish to use scales of factors other than these. On the other hand, the investigator may often need only the scales of one factor, most likely the evaluative factor. This would be true in studies of attitudes and values. In some research too, an investigator may decide it is necessary to include scales whose factorial composition is not known.

The second criterion of scale selection, relevance to the concepts used, is more difficult to satisfy. In many cases, it is obvious that an adjective pair is related to a concept. However, according to Osgood, scales of unknown factorial composition are often relevant to certain problems, and when using such untried scales, one should attempt to determine the factorial identity of the

scales. That is, certain adjective pairs may seem irrelevant to the concepts judged, but one cannot always be certain of relevance. Meanings are rich and complex, and an apparently irrelevant adjective pair may turn out to be relevant.

There are three main sources of variance: concepts, scales, and subjects (and, of course, error). Kerlinger (1964) discussed analytic techniques that can be used for the SD in the following way:

. . .the scores can be analyzed for differences between concepts, between scales, between subjects, or any combination thereof. However there are ways of reducing the data to two categories, usually concepts and scales, or concepts and factors. The SD data are unique in that the data of one individual can be analyzed, as well as the data of groups of individuals [p. 572].

According to Osgood the most obvious analysis of the data would be to compare the means of the concepts. There is a question of the legitimacy of the use of the usual statistical tests of significance with the data from one individual. However, if the scores were averages, e.g., means or medians - of the scores of a group of individuals, it would be possible to use such tests.

Analysis of Data

Data obtained from the students in response to the first part of the semantic differential, entitled ORAL EXAMINATION, were factor-analyzed and the unrotated factor matrix (see Appendix E) was produced. The varimax rotation procedure (see Appendix F) revealed three principal factors or dimensions which the researcher called EVALUATION, ATMOSPHERE, and POTENCY. Scales (or adjective pairs) that had loadings

greater than 0.5 on each of these three factors, were grouped together. Thus, of the 26 scales used to respond to the concept, ORAL EXAMINATION, there were 16, 9, and 6 scales that belonged to the EVALUATION, ATMOSPHERE and POTENCY factors respectively. There were some scales that satisfied the 0.5 loading norm on two factors. Such scales were included in both factors.

For each scale, the scores of every individual in his respective group (i.e., the FB or NO-FB group) were added and the mean value was calculated. Hence it was possible to test the significance of the difference between the two groups on every scale by using Student's t-test. This procedure permitted the researcher to examine the analysis of the responses made to individual scales.

In order to look at each factor as a whole, it was necessary to combine the scores of every individual on the scales belonging to the same factor and to calculate the means for the FB and NO-FB groups. As discussed previously, the EVALUATION factor corresponds to a measure of attitude. It was therefore possible to compare the attitude of the two groups and to test the significance of the difference between the means by using Student's t-test.

The factor analysis of the concept, ORAL EXAMINATION also revealed the scales belonging to the factor, ATMOSPHERE. Since a separate section of the SD (see part iv, p. 32) had also utilized the concept, ATMOSPHERE DURING THE ORAL EXAMINATION, as a stimulus, it was possible to check the reliability of the responses made by the students. If the responses in one section of the SD dealing with atmosphere (a factor obtained by factor-analyzing the concept, oral examination) were signi-

ificantly different from those in another section also concerned with atmosphere, then the reliability of the SD as a measuring instrument, would be questioned.

Scales in Parts ii and iii of the SD concerned with the oral examination as a learning experience and as a means of evaluating students were similarly combined. The t-test was used to compare the difference between the means of the FB and NO-FB groups.

An analysis, again using the t-test, was carried out on the responses made on the questionnaire in order to test the significance of the differences between the means of the FB and NO-FB groups. The t-test was selected as the most appropriate statistical procedure since the main assumption of homogeneity of variance (see Appendix G) was satisfied.

Reliability of grading in the oral examination was determined by calculating the Pearson Product Moment Coefficients of Correlation between the grades assigned by each examiner. These correlation coefficients were calculated for each grading variable as well as for the composite grade. A two-way analysis of variance was done to compare the marks assigned by examiners and students for feedback and no-feedback groups.

CHAPTER V

ANALYSIS AND INTERPRETATION OF RESULTS

Reliability of the Oral Examination

A review of previous attempts (Barnes and Pressey, 1929; Trimble, 1934; Carter, 1962) to measure the reliability of oral examinations has revealed low correlations between the grades or rank orders assigned by different examiners. In looking at the procedures utilized in many of these studies, the writer has noted the following weaknesses that may have led, in part, to the conclusion that the oral examination is too subjective a tool to provide consistent grading:

- (a) Criteria for grading were not properly defined,
- (b) Criteria for grading were not comprehensive enough,
- (c) The amount of time spent in examining students was insufficient to provide an adequate evaluation, and
- (d) Correlations were made between marks assigned in the oral and marks attained on written examinations. The assumption that they are measuring the same aspects of student ability is not tenable.

In this study, an effort was made to define the grading criteria (see Appendix H) so that they properly reflected the aims of evaluation. A great deal of consideration was also given to deciding what criteria were to be used for grading. In addition it was the opinion both of the examiners and of the majority of students that sufficient time (about thirty minutes) was allotted for each oral.

For each criterion variable, correlation coefficients were calculated between the grades assigned by each of the examiners as well as by the examinees. A similar analysis was carried out on the composite grades. A summary of these results is presented in Table 1.

TABLE 1

PEARSON PRODUCT MOMENT COEFFICIENTS OF CORRELATION OF GRADES
ASSIGNED BY EXAMINERS AND STUDENTS ON THE ORAL EXAMINATION

Raters Compared	Criterion Grading Variables					Composite Ratings
	Ability to Report Facts Clearly	Understanding of Concepts in Essay	Ability to Present Summary of Essay	Ability to Argue Topics Reported in Essay	Ability to Discuss Rele- vance of Topic for Education	
1 & 2	.6158	.5369	.4475	.4710	.4384	.6946
1 & 3	.4654	.4873	.5113	.4681	.5160	.6021
2 & 3	.4982	.4945	.4624	.4664	.2904	.6083
1 & Ss	.1553	.2361	.1742	.2245	.0583	.2583
2 & Ss	.1300	.1365	.3119	.2493	.0347	.2006
3 & Ss	.3370	.5675	.2892	.2348	.1899	.4911

As seen in Table 1, the correlation coefficients between each pair of examiners for the composite grades were much higher (all were above .60) than between examiners and students. This difference was probably a function of (a) the greater experience of the examiners in comparison to the examinees with regard to the grading procedures, (b) the greater opportunity the examiners had for comparing student performance - the examinee did not have a chance to measure his performance against that of the other students and (c) the consensus arrived at by the examiners in deciding the exact nature of the grading criteria.

The results also showed that for the examiners the correlation coefficients obtained for the composite ratings were higher than that obtained for the individual criterion variables. That is, there was better statistical agreement between the examiners for the total grade than for the grades assigned on individual variables.

Preamble to Presentation of Results

In attempting to arrive at conclusions regarding the hypotheses, the researcher made use of a questionnaire as well as a semantic differential. Data obtained from responses made on these two instruments were very often used to supplement each other. In looking at the analyses of the responses to the questionnaire, one should be aware of the following:

- (a) The range on this instrument was from 0 to 4, with the lower end of the scale reflecting negative responses.
- (b) The level of significance has been set at the .05 level.
- (c) Even though the analyses may reveal differences significant at the .05 level of probability or less, it is necessary to take into account both the magnitude and direction of the responses.

With regard to the analysis of the data obtained on the semantic differential, one should also be aware of the following points:

- (a) A seven-point scale (1 to 7) was used with the lower end reflecting the more positive aspect of the dichotomy, i.e., it is a response measuring both magnitude and direction.
- (b) Scales (or adjective pairs) belonging to the same dimension were examined individually and were also combined and the difference between the means of the two groups (FB and NO-FB) was tested.
- (c) While the level of significance has been set at the .05 level, any conclusions made must take into account the direction of response of the two groups. For example, if the mean of one group was 3.82 and that of the other 2.50, this difference might be significant at the .05 level. However the direction of response of both groups indicates favourable responses.

Results

A summary of the comparisons of means for the feedback and no-feedback groups on the questionnaire is presented in Table 2. As noted previously, the questions have been grouped according to the following major dimensions of the study:

- (a) Oral examination as a learning experience.
- (b) Oral examination as a means of evaluation.
- (c) Anxiety and the oral examination.
- (d) Atmosphere during the oral.
- (e) Communicative ability and the oral examination.
- (f) General attitude to the oral examination.

(a) Oral Examination as a Learning Experience

The following hypotheses pertaining to the oral examination as a learning experience were advanced. In comparison to students in the no-feedback group, students in the feedback group would

TABLE 2

SUMMARY OF COMPARISONS OF MEANS FOR FEEDBACK
AND NO-FEEDBACK GROUPS ON THE QUESTIONNAIRE

(a) Questions pertaining to the oral examination as a learning experience.	Combined Means		t-Ratio	Level of Significance
	Feedback Group	No-Feedback Group		
12. To what extent would you view the O.E. as a learning experience?	2.30	1.82	1.40	.08
1. To what extent do you feel you have gained in terms of knowledge as a result of having the O.E.	1.90	.82	3.46	.0006
2. To what extent do you feel you have gained in terms of better understanding of the implications of your essay for teaching?	2.30	1.45	2.41	.01
(b) Questions pertaining to the oral examination as a means of evaluation.				
13. To what extent would you view the O.E. as a better evaluative tool than a written examination?	2.45	1.68	2.035	.02

(b) Continued	Combined Means		t-Ratio	Level of Significance
	Feedback Group	No-Feedback Group		
16. To what extent do you feel the O.E. provided you with an opportunity to discuss or defend your essay?	2.55	1.82	1.930	.03
5. Did the O.E. give you an opportunity to explain more than you would have normally communicated in a written examination?	2.30	1.41	2.325	.01
(c) Questions pertaining to anxiety and the oral examination				
6. How would you rate your anxiety level just before the O.E.? (no anxiety - great anxiety)	2.32	2.82	-1.609	.06
7. How would you rate your anxiety level during the O.E.? (no anxiety - great anxiety)	2.20	2.77	-1.787	.04
8. To what extent do you feel that your anxiety adversely affected your performance during the O.E.? (no adverse effects - great adverse effects).	1.75	2.64	-2.496	.008

(d) Questions pertaining to atmosphere during the oral examination	Combined Means		t-Ratio	Level of Significance
	Feedback Group	No-Feedback Group		
3. How would you describe the atmosphere created during the O.E.?	3.15	1.64	4.802	.0001
4. Do you feel that the atmosphere during the O.E. (made no difference in your performance - made a difference in your performance)?	2.95	2.77	0.795	.215
(e) Questions pertaining to communicative ability and the oral examination				
14. How do you rate yourself on your ability to communicate verbally?	2.15	2.27	- .484	.315
15. To what extent do you feel that your ability to communicate verbally assisted you in the O.E.?	2.10	1.73	1.077	.144

Hypothesis 1: view the oral examination as more of a learning experience.

Hypothesis 2: experience greater gain in terms of knowledge as a result of having the oral examination.

Hypothesis 3: feel that they have gained more in terms of better understanding of the implications of their essays for teaching.

A t-test of the responses made to Question 12 (Table 2) revealed that students who were given feedback ($\bar{x} = 2.30$) viewed the oral examination as more of a learning experience than those in the no-feedback group ($\bar{x} = 1.82$). However the difference between these means was not significant ($p = .08$). Hypothesis 1 was not supported by analysis of the results from the questionnaire.

Analysis of the data elicited on the semantic differential regarding the concept, 'The Educ. Psy. 471 oral examination as a learning experience' demonstrated significant differences between the two groups on all the scales except the first (see Appendix I for a t-test of the differences between the means of each scale for the two groups). A profile of the responses made to these scales is presented in Figure 1.

In contrast to the NO-FB group, students in the FB group felt that as a learning experience, the oral was more helpful, valuable, successful, good and significant. The difference between the two groups on this question was a significant one.

In order to examine the concept as a whole, all the scales were combined. A summary of the comparison of the means for the FB and NO-FB groups is presented in Table 3. A highly significant difference ($p = .004$) between the two groups was found. The mean value for the NO-FB group was 4.44 on the seven-point scale while that for the FB group was 3.32. Since a

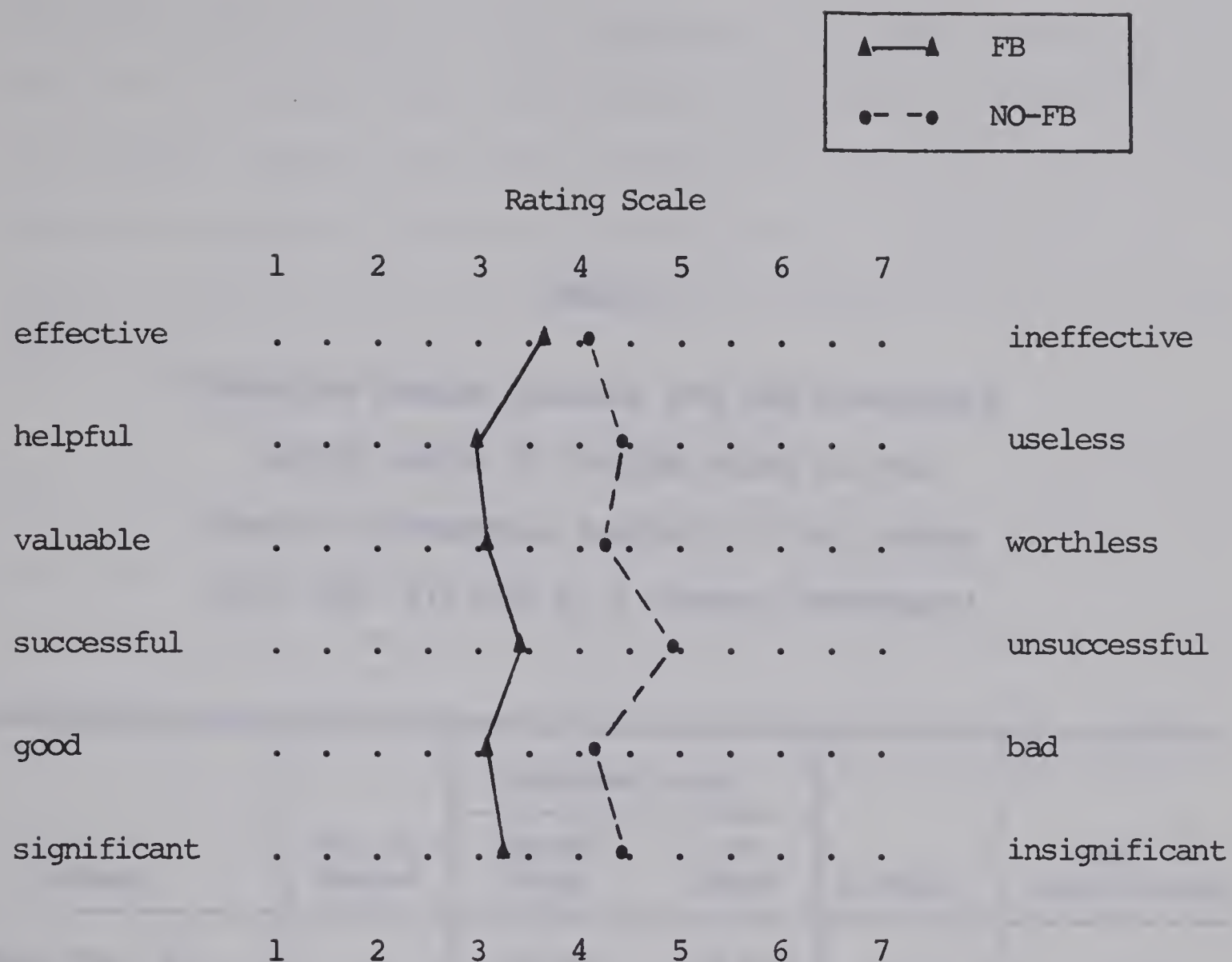


Figure 1 Profile of Semantic Differential Responses to the Concept, 'Educ. Psy. 471 Oral as a Learning Experience' for Feedback and No-Feedback Groups

TABLE 3

COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK
(NO-FB) GROUPS OF COMBINED MEANS FOR THE
SEMANTIC DIFFERENTIAL RESPONSES TO THE CONCEPT
'EDUC. PSY. 471 ORAL AS A LEARNING EXPERIENCE'

Concept	No. of Scales	Combined Mean		t-Ratio	Level of Significance
		NO-FB Group	FB Group		
Educ. Psy. 471 Oral as a learn- ing experience	6	26.64 (4.44)*	19.90 (3.32)*	2.74	.004

* Equivalent value on a seven-point scale obtained by dividing the combined mean by the number of scales.

score lower than 4 on this measure indicates a favourable response, it can therefore be concluded that the FB group had favourable responses toward the oral as a learning experience in comparison to the NO-FB group and the difference between the two groups was significant. Hypothesis 1 was supported by the data elicited on the semantic differential but not by the data from the questionnaire.

In terms of a reported gain in knowledge as a result of having the oral examination (see Question 1) the difference between the means of the two groups was highly significant ($p < .001$) with the FB group ($\bar{x} = 1.90$) reporting a greater increase in knowledge than the NO-FB group ($\bar{x} = .82$). Hypothesis 2 was supported. It should be noted, however, that the means of both groups on this question did not exceed 2, the mid-point on the scale, 0 (no gain) to 4 (very much gain).

In terms of better understanding of the implications of their essays for teaching (see Question 2) a significant difference ($p = .01$) was found between the two groups, with the FB group ($\bar{x} = 2.30$) again reporting more gain in comparison to the NO-FB group ($\bar{x} = 1.45$). Hypothesis 3 was supported.

(b) Oral Examination as a Means of Evaluation

The following hypotheses pertaining to the oral examination as a means of evaluation were advanced: In comparison to students in the no-feedback group, students in the feedback group would

Hypothesis 4: rate the oral examination as a better evaluative tool than a written examination.

Hypothesis 5: feel that the oral examination provided them with a greater opportunity to discuss or defend their essays.

Hypothesis 6: feel that the oral examination gave them an opportunity to communicate more than they would have normally communicated in a written examination.

Both groups viewed the oral examination as a better evaluative tool than a written examination (see Question 13, Table 2), with the FB group ($\bar{x} = 2.45$) showing a more positive response than the NO-FB group ($\bar{x} = 1.68$). This difference was significant at the .02 level. Hypothesis 4 was supported. In an attempt to delineate some of the ways in which the oral was considered to be superior to a written examination, questions 16 and 5 were included in the study.

On the question of the extent to which the oral provided an opportunity to discuss or defend the essay (see Question 16), it was found that the difference between the means of the two groups was significant ($p = .03$): the FB group ($\bar{x} = 2.55$) felt that the oral provided them with a greater opportunity to discuss or defend their essay than the NO-FB group ($\bar{x} = 1.82$). Hypothesis 5 was supported.

A significant difference ($p = .01$) was also noted between the two groups on the question, 'Did the oral give you an opportunity to explain more than you would have normally communicated in a written examination?' with the direction of significance being FB group > NO-FB group. Hypothesis 6 was supported.

In summary it should be pointed out that both groups were favourable to the oral as an evaluative procedure, only the FB group was significantly more so.

Further support for this conclusion was obtained from data derived from the semantic differential. Of the 9 scales used to respond to the concept, 'The oral as a means for evaluating students' only responses to

4 (flexible-rigid, like-dislike, valuable-invaluable, good-bad) showed differences that were significant between the two groups (see Appendix J). However it should be noted that for the FB group, the means for all 9 scales were less than 4 which means that the responses were favourable; as well the means for 3 scales were also below 4 for the NO-FB group. A profile of these responses to the scales is shown below in Figure 2.

An overall assessment of the oral as an evaluative tool was then made by combining the above scales. A significant difference ($p = .01$) was found between the two groups (see Table 4). The mean for the NO-FB group was 4.14 and that of the FB group, 3.42.

On the whole then, the analysis of the data from the semantic differential supported the conclusion previously reached that the FB group, in comparison to the NO-FB group, viewed the oral as a better evaluative tool than a written examination.

(c) Anxiety and the Oral Examination

The following hypotheses concerning anxiety and performance as it pertains to the oral examination were advanced: In comparison to students in the feedback group, students in the no-feedback group would

Hypothesis 7: show no difference in their reported level of anxiety just before the oral examination.

Hypothesis 8: rate themselves as being more anxious during the oral examination.

Hypothesis 9: feel that their anxiety adversely affected their performance during the oral examination.

As the results of the comparison of means for Question 6 (Table 2) has shown, both FB and NO-FB groups experienced some anxiety just before the oral. But the difference between the means of the two groups was only

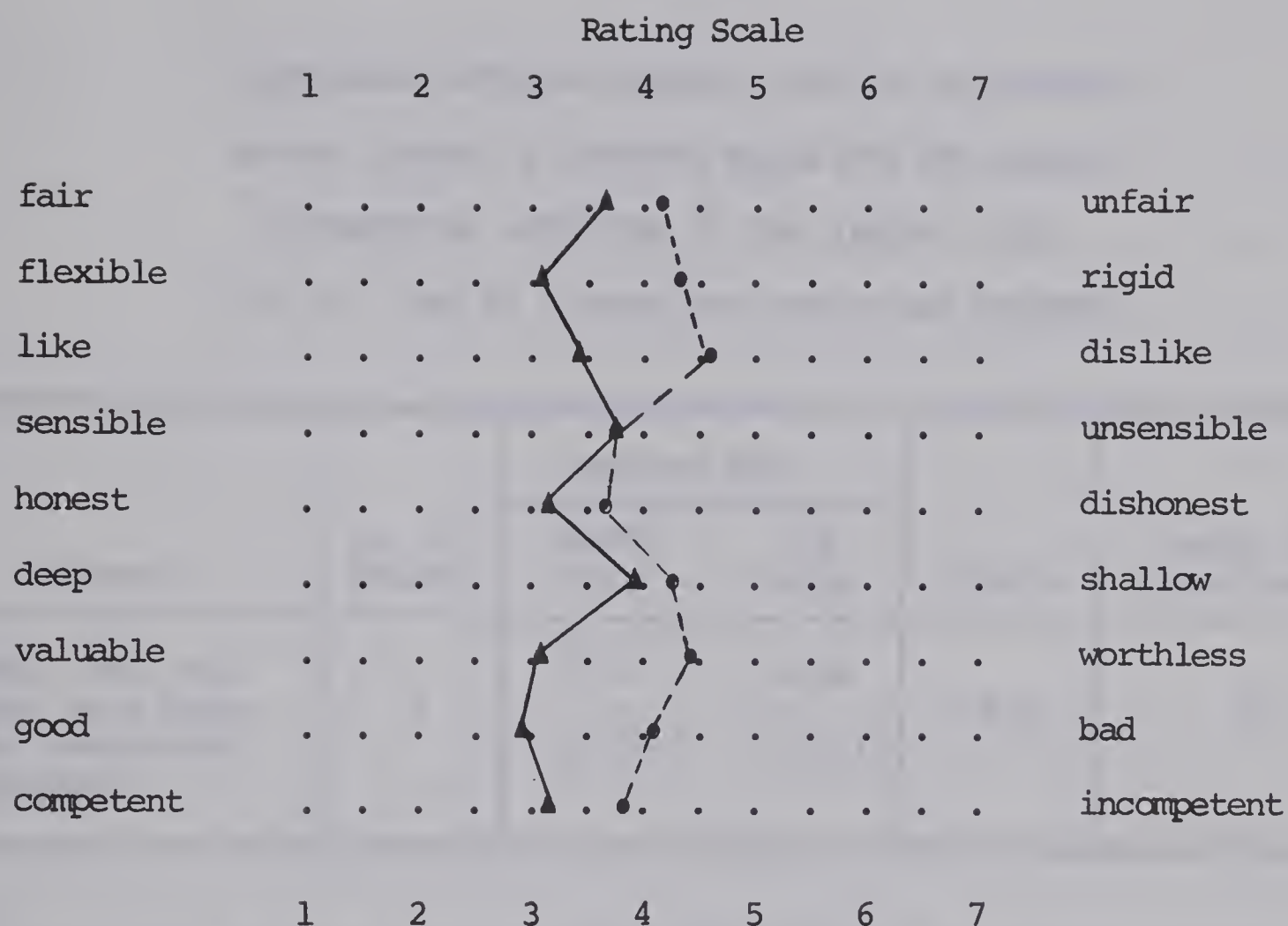
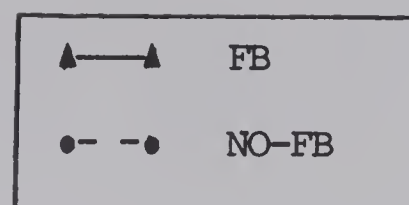


Figure 2 Profile of the Semantic Differential Responses to the Concept, 'Educ. Psy. 471 Oral as a Means for Evaluating Students' for Feedback and No-Feedback Groups

TABLE 4

COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK
(NO-FB) GROUPS OF COMBINED MEANS FOR THE SEMANTIC
DIFFERENTIAL RESPONSES TO THE CONCEPT 'EDUC.
PSY. 471 ORAL AS A MEANS FOR EVALUATING STUDENTS'

Concept	No. of Scales	Combined Mean		t-Ratio	Level of Significance
		NO-FB Group	FB Group		
Educ. Psy. 471 Oral as a means for evaluating students	9	37.27 (4.14)*	30.80 (3.42)*	2.133	.01

* Equivalent value on a seven-point scale obtained by dividing the combined mean by the number of scales.

approaching significance ($p = .06$) with the FB group ($\bar{x} = 2.32$) indicating somewhat less anxiety than the NO-FB group ($\bar{x} = 2.82$). Hypothesis 7 (a null hypothesis) was supported.

During the oral, the reported level of anxiety of both groups was slightly reduced in comparison to before the oral: this time the difference between the groups was significant ($p = .04$). The means for the FB and NO-FB groups were 2.20 and 2.77 respectively. Hypothesis 8 was supported.

The analysis also revealed (Question 8, Table 2) that the NO-FB group ($\bar{x} = 2.64$) felt that their anxiety adversely affected their performance during the oral; however, the FB group ($\bar{x} = 1.75$) indicated somewhat less adverse effects on performance as a result of their anxiety. The difference between the groups on this question was highly significant ($p = .008$). Hypothesis 9 was supported.

It is relevant at this point to compare the actual marks of the students (as graded by the examiners) with their self-scored marks (as graded by the students themselves) for both FB and NO-FB groups. The results of this comparison is presented in Table 5.

As seen in Table 5, the mean of the marks assigned by the examiners to the NO-FB group ($\bar{x} = 16.53$) was slightly higher than that of the FB group ($\bar{x} = 15.92$). A similar trend was evident for the marks assigned by the students in the two groups; the means for the NO-FB and FB groups were 15.87 and 15.57 respectively. In addition, the marks assigned by the examiners to both groups were slightly higher than the marks the students assigned to themselves.

A two-way analysis of variance was done to test the main effects of

TABLE 5

MEANS OF MARKS ASSIGNED BY EXAMINERS AND STUDENTS
FOR FEEDBACK AND NO-FEEDBACK GROUPS

Group	n	Marks* Assigned by	
		Examiners	Students
Feedback	21	15.92	15.57
No-Feedback	22	16.53	15.82

* Marks were based on a 25-point total score with 5 points assigned to each of the 5 grading variables.

factor A (FB and NO-FB groups) and factor B (marks assigned by examiners and students) with factor B being a repeated measure. A summary of the analysis is presented in Table 6. It was found that there were no significant differences for the main effects or their interaction. That is, the marks assigned by the examiners to the FB and NO-FB groups were not significantly different from those assigned by the students themselves. In addition, students in the FB group did not assign to themselves marks that were significantly different from the marks assigned by the students in the NO-FB group.

In summary, then, both groups felt that their performances were adversely affected by their anxiety but there was a significant difference between the two groups; the NO-FB group reported significantly greater adverse effects on performance than the FB group.

(d) Atmosphere During the Oral Examination

The following hypotheses relating to atmosphere during the oral examination were advanced: In comparison to students in the feedback group, students in the no-feedback group would

Hypothesis 10: perceive the atmosphere during the oral examination to be unfriendly, threatening, impersonal, etc.

Hypothesis 11: feel that the atmosphere during the oral examination made more of a difference in their performance.

Analysis of responses made to Question 3 (Table 2) demonstrated that the difference between the means for the FB and NO-FB group relating to atmosphere during the oral, was highly significant ($p < .001$); the NO-FB perceived the atmosphere as being somewhat impersonal ($\bar{x} = 1.64$) while the FB group felt that it was very friendly ($\bar{x} = 3.15$). Additional support for this interpretation was also evident from data obtained on the semantic

TABLE 6

ANALYSIS OF VARIANCE OF MARKS CLASSIFIED BY EXAMINERS AND STUDENTS

(FACTOR B) FOR FEEDBACK AND NO-FEEDBACK GROUPS (FACTOR A)

Source of Variation	SS	DF	MS	F	P
Between Subjects	449.12	42			
'A' Main Effects	3.96	1	3.96	0.37	.55
Subjects Within Groups	445.15	41	10.86		
Within Subjects	427.03	43			
'B' Main Effects	0.79	1	0.79	0.08	.78
'A * B' Interaction	0.35	1	0.35	0.03	.85
'B' x Subj. Within Groups	425.88	41	10.39		

differential. Responses on 7 of the 11 scales concerned with the concept, 'Atmosphere during the oral' were significantly different between the two groups (Appendix K). A profile of these responses is shown in Figure 3.

When the above scales were combined in order to examine the overall response to the concept, 'Atmosphere during the oral', a highly significant differences was also found between the two groups. These results are presented in Table 7.

An interesting observation is the fact that the data derived from the semantic differential which had as its heading, ORAL EXAMINATION (and which was factor-analyzed to separate out the scales that had high loadings on the dimension, ATMOSPHERE, (see Appendix F) was consistent with that obtained from looking at the concept, ATMOSPHERE DURING THE ORAL, separately (see Table 7). In both cases the level of significance was at the .001 level.

A further finding related to the atmosphere during the oral was that the difference between the two groups on the question of whether the atmosphere affected their performance was not significant (see Question 4, Table 2). The mean for the NO-FB group was 2.77 and that for the FB group was 2.95. Hypothesis 11 was rejected. However it is important to note that the direction of response for the two groups was similar, i.e., both groups felt that the atmosphere during the oral made a tremendous difference in their performance. However, the marks assigned by the examiners and students to the two groups were not significantly different.

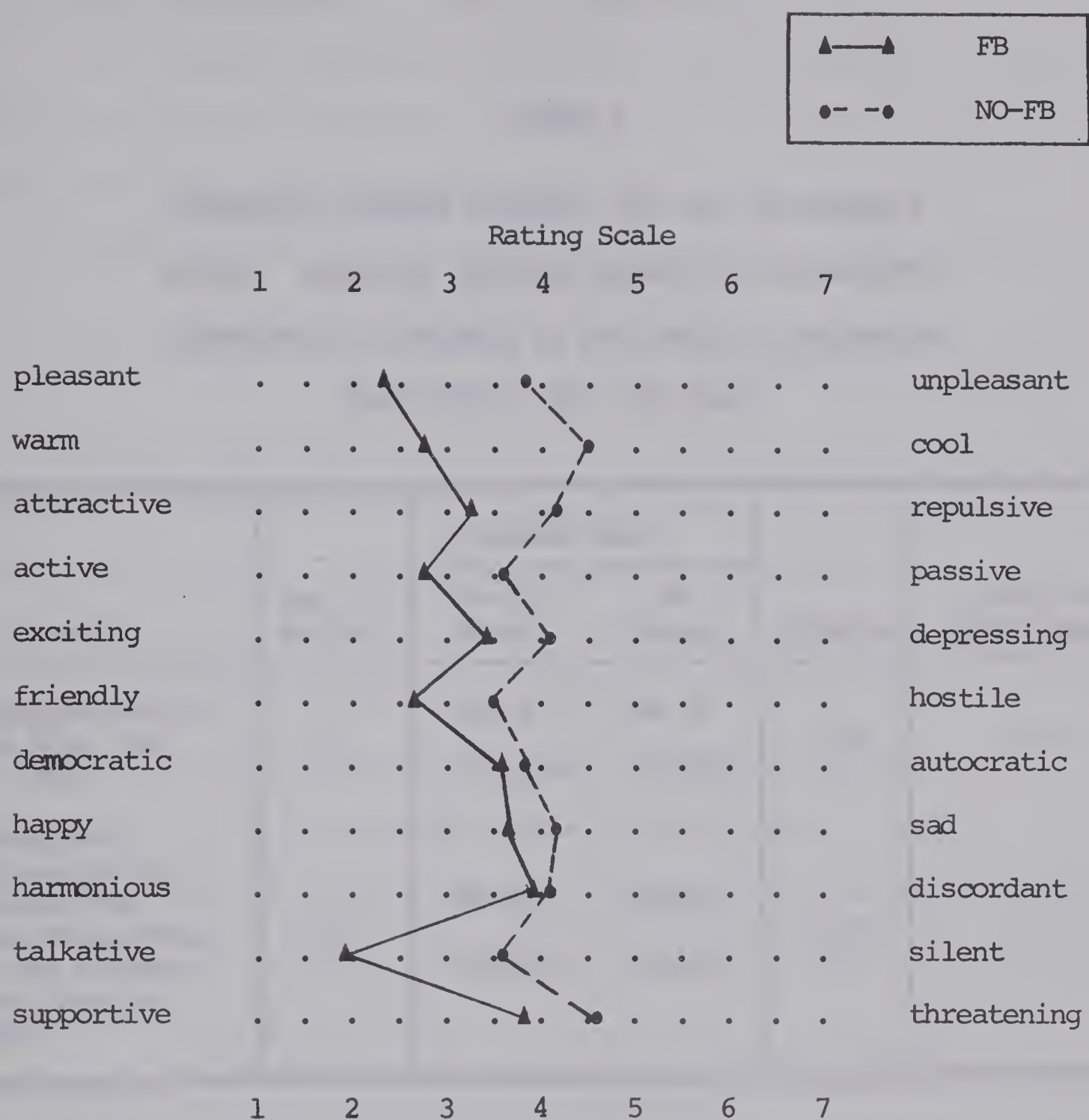


Figure 3 Profile of the Semantic Differential Responses to the Concept 'Atmosphere During the Oral' for Feedback and No-Feedback Groups

TABLE 7

COMPARISON BETWEEN FEEDBACK (FB) AND NO-FEEDBACK
(NO-FB) GROUPS OF COMBINED MEANS FOR THE SEMANTIC
DIFFERENTIAL RESPONSES TO THE CONCEPT 'ATMOSPHERE
DURING EDUC. PSY. 471 ORAL'

Concept	No. of Scales	Combined Mean		t-Ratio	Level of Significance
		NO-FB Group	FB Group		
Atmosphere during Educ. Psy. 471 oral	11	44.14 (4.01)*	34.35 (3.12)*	3.225	.001
Atmosphere (Dimension obtained from factor-analysis of the concept Oral Examination)	10	40.95 (4.09)*	30.65 (3.06)*	3.065	.001

(e) Communicative Ability and the Oral Examination

The following hypotheses concerning communicative ability and the oral examination were advanced: In comparison to students in the no-feedback group, students in the feedback group would

Hypothesis 12: rate themselves no differently in their ability to communicate verbally.

Hypothesis 13: feel that their ability to communicate verbally assisted them more during the oral examination.

As predicted, there was no significant difference between the two groups in their ratings of their ability to communicate verbally (see Question 14, Table 2). The mean value for the NO-FB group ($\bar{x} = 2.27$) was only slightly higher than that of the FB group ($\bar{x} = 2.15$). Hypothesis 12 (which is a null hypothesis) was accepted.

However, the hypothesis (13) that students in the FB group, in comparison to those in the NO-FB group, would feel that their ability to communicate verbally assisted them more in the oral was rejected (Question 15, Table 2). The FB group ($\bar{x} = 2.10$) showed only slightly higher ratings than the NO-FB group ($\bar{x} = 1.73$).

(f) General Attitude Toward the Oral Examination

The following hypothesis concerning general attitude toward the oral examination was advanced:

Hypothesis 14: compared to students in the no-feedback group, students in the feedback group would have more positive attitudinal responses toward the oral examination.

As discussed previously, the EVALUATIVE dimension of the semantic differential corresponds to a measure of attitude. This dimension was determined by factor-analysis of the responses made to the general concept

ORAL EXAMINATION. Factor-analysis provides a mathematical model in which a series of scores are intercorrelated to determine the number of dimensions the test space occupies, and to identify these dimensions in terms of traits or other general concepts. By analyzing the pattern of inter-correlations, the operation of one or more underlying traits or other sources of common variance is inferred. For the semantic differential, the identification of clusters of adjectives was made easier by rotating the axes of the semantic space in order to maximize the sources of common variance. In this way, from the factor-analysis of the semantic differential responses to the concept, ORAL EXAMINATION, it was possible to separate out the evaluation, atmosphere, and potency dimensions. The scales that were collected under the evaluation dimension were the ones that had a minimum factor loading of approximately 0.50 (varimax rotation, see Appendix F). A total of 16 scales satisfying the above criterion were found using this procedure and a profile of the responses of the FB and NO-FB groups is presented in Figure 4.

Comparison of the two groups revealed differences in means that were significant beyond the .05 level on the following scales: fair-unfair, like-dislike, effective-ineffective, helpful-useless, good-bad, attractive-repulsive, successful-unsuccessful, valuable-worthless, accepting-rejecting, involved-withdrawn, and significant-insignificant (see Appendix L).

On all 16 scales, means of ratings by the FB group were less than 4, thus indicating positive responses. However, the NO-FB group gave ratings of less than 4 on only 7 of the 16 scales.

In order to get at the total attitudinal response of the two groups it was necessary to combine the 16 scales. A summary of the

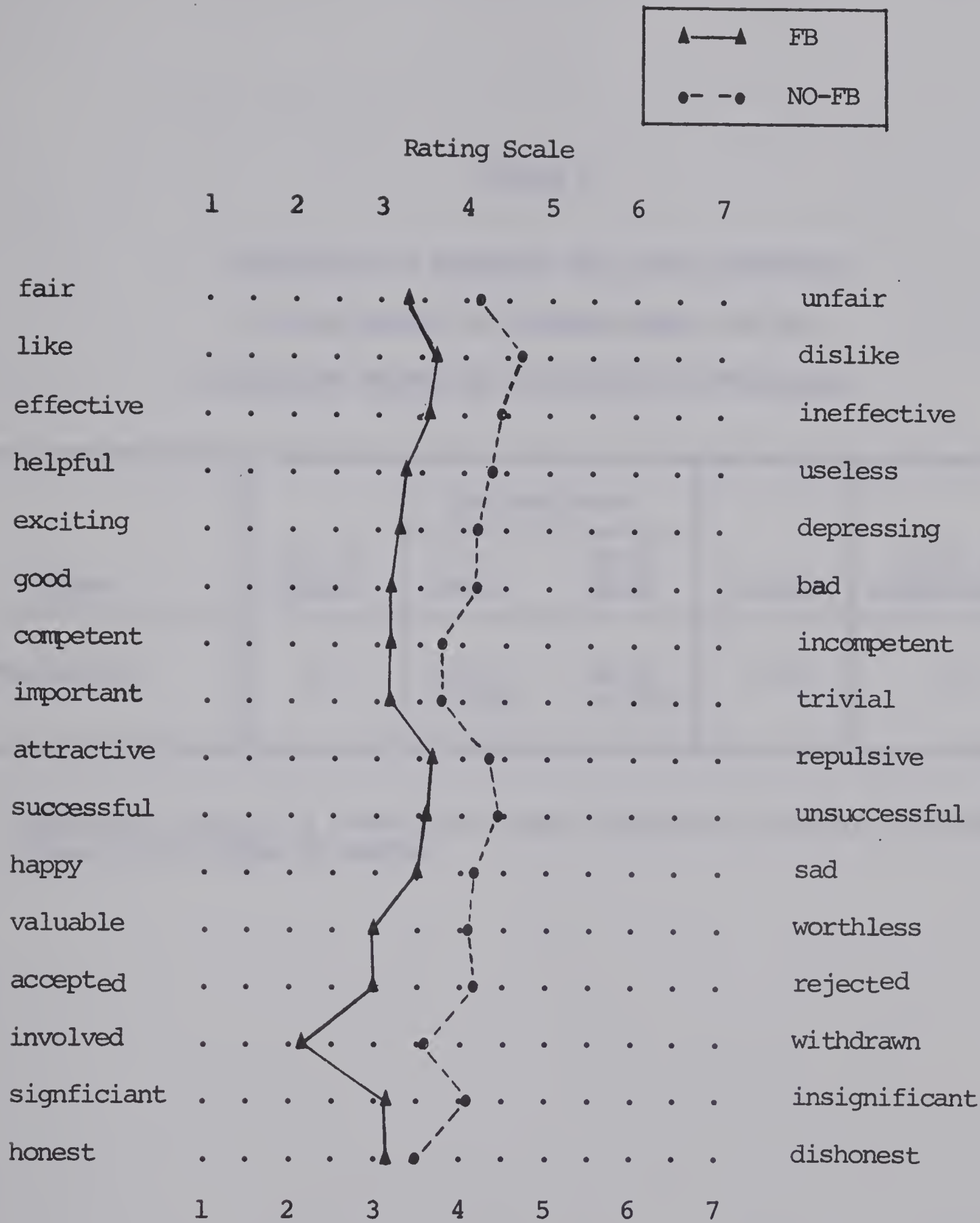


Figure 4 Profile of Semantic Differential Responses made by Feedback and No-Feedback Groups to the scales belonging to the Evaluative Dimension

TABLE 8

COMPARISON OF FEEDBACK (FB) AND NO-FEEDBACK
(NO-FB) GROUPS OF COMBINED MEANS FOR THE
EVALUATION FACTOR OF THE SEMANTIC DIFFERENTIAL

Factor	No. of Scales	Combined Mean		t-Ratio	Level of Significance
		FB Group	NO-FB Group		
Evaluative	16	52.15 (3.36)*	66.00 (4.13)*	-2.26	.01

* Equivalent value on a seven-point scale obtained by dividing the combined value by the number of scales.

analysis of the combined ratings for the evaluative dimension is presented in Table 8.

The results from Table 8 revealed that the FB group ($\bar{x} = 52.15$) on the whole, had a significantly more positive attitudinal response towards the oral than the NO-FB group ($\bar{x} = 66.00$). The equivalent value on a seven-point scale for the FB group was 3.26 and that for the NO-FB group was 4.13. (Note: a score less than 4 indicates a favourable response and vice-versa.) Hypothesis 14 was supported.

CHAPTER VI

DISCUSSION

Examinations serve to evaluate a student's knowledge in a particular area. The oral has an additional dimension in that it provides a unique opportunity for examiners and examinees to communicate with each other and hopefully, to help the student gain insights into differing views of the subject under consideration. To the extent that this interaction is meaningful, it becomes a learning experience for the student. The oral also provides the examiners, with, perhaps, a more comprehensive look at the student's ability to conceptualize as well as to defend and discuss his views. In situations where factual knowledge is to be tested, the use of an oral examination may be unnecessary; but where there is need to go beyond mere facts and to examine the implications arising from certain assumptions made or viewpoints adopted by the student, the oral becomes an invaluable tool for this purpose.

It is the writer's contention that the amount and type of feedback provided to the examinee are crucial variables in determining whether he perceives the oral to be a learning experience. While it might be argued that knowledge of results of written examinations is, in effect, feedback, this procedure is nevertheless often deficient in two ways: (a) feedback is generally given in the form of a grade for the entire examination and hence does not provide any meaningful analysis of performance, and (b) immediate feedback on specific content is not available and within the interpretation of learning theory

tradition, this means that its effectiveness is reduced.

Another issue that arises is related to the reliability of a student's perception of how much or how little he has gained. In other words, has the student learned as much or as little as he claims? Admittedly it becomes a matter of subjective evaluation on his part. Besides a lot of incidental and unconscious learning can take place. On the other hand there is no reason to reject the student's assertion that he has understood and assimilated a particular concept. Moreover, if a student feels that he has gained in knowledge and understanding, it cannot help but affect his entire perception of the subject matter in a positive manner and undoubtedly could influence the effectiveness of subsequent learning.

The results obtained from the analysis of the responses made by groups having differing amounts of feedback are fairly conclusive: the FB group is evidently more disposed toward the oral as a learning experience than the NO-FB group. It is interesting to note that on this particular issue, analysis of responses made on the semantic differential revealed significant differences overall between the groups but failed to show substantial differences when the data from the questionnaire was processed. It may be that as an instrument the semantic differential is more refined at teasing out differences than the questionnaire. In addition, the rather arbitrary nature of the scale used on the questionnaire may be inadequate in providing enough response alternatives. Nevertheless, as the results have shown, subsequent analysis of responses to questions relating to a reported gain in knowledge and understanding of implications, has

provided evidence of significant differences between the groups, in favour of the FB group.

In any examination, the process and procedures by which a student is evaluated should be closely linked to the objectives of a program of study, and, of course, to the actual work that was covered over a period of time. A more technically oriented course would, of necessity, require proficiency and factual knowledge in certain key sections of the program. But in many other areas of educational endeavour, the ever-changing needs of the student within an increasingly complex milieu demand a more comprehensive approach to the problems at hand. Consistent with this view is the notion that evaluative tools must be flexible and thorough enough to give a more valid assessment of how far the student has progressed. Written examinations can often fulfill the above objective up to a point. But it is to oral examinations that one must turn in order to determine the full extent of the degree and kind of assimilation that has taken place. Some examiners may feel that their own judgment of how well an examination fulfills its function is sufficient, neglecting to tap student opinion.

This study has provided students with an opportunity to compare the oral with written examinations. It was found that both FB and NO-FB groups were more favorable towards the oral as an evaluative procedure than towards a written examination, the FB group more so than the NO-FB group. The two groups also felt that the oral afforded them an opportunity to discuss or defend their essays, in addition to giving them an opportunity to explain more than they would have normally

communicated in a written examination. Again the FB group reported response more positive than those of the NO-FB group.

These results support the contention that, in the students' view, the oral, even without feedback, can get at a more adequate assessment of students, in comparison to a written examination.

It was initially hypothesized that the anxiety level of both the FB and NO-FB groups just before the oral would not be significantly different since the students were randomly selected for each treatment condition. However the results showed the difference in anxiety level was approaching significance. This difference might be the result of a post-treatment 'halo' effect; that is, students were asked to rate their pre-treatment anxiety level after they had the oral. For students in the FB group, because of the positive nature of the interaction with the examiners and the feedback they received, they probably felt a reduction in anxiety after they had the oral, and hence in reflecting on their pre-treatment anxiety, would tend to rate it a little lower than it actually was. The exact opposite effect may have occurred for students in the NO-FB condition. A defense of the research strategy concerned with the measurement of anxiety is necessary at this point. The writer decided against the administration of an anxiety questionnaire to each student immediately before the oral examination for the following reasons: (a) it would probably have alerted some students to the fact that a research project was in progress, (b) it was likely to unnecessarily raise the anxiety level of the students, and (c) there is no reason to believe that the administration of an anxiety questionnaire at that time would

result in a more accurate assessment of anxiety.

One of the crucial factors that examinees are invariably concerned with is anxiety and its effect on performance. Although it is difficult to eliminate anxiety in examinations, especially oral examinations (and it is questionable whether it is desirable to do so) examiners should attempt to limit the potentially adverse effects that too much anxiety may produce. The most significant finding in this part of the study concerning anxiety, is that there is a highly significant difference between the FB and NO-FB groups in terms of their perception of the effect of anxiety on performance. Yet a comparison of the grades assigned by examiners and students demonstrated no significant difference between the groups. This is not to say that anxiety had no influence on performance. It simply means that as a variable, the effects of self-reported anxiety may have been overestimated.

The hypotheses in the section of the study involving atmosphere during the oral examination, were included for two reasons: (1) to determine whether the examiners were successful in differentiating between the two treatment groups and creating two atmospheres and (2) to find out whether the examinees perceived any relationship between the atmosphere during the oral and their performance.

One assumption made was that the atmosphere during the oral was indirectly related to the amount of feedback given. It was expected that students who received feedback would feel the atmosphere during the oral was friendly, warm, pleasant, supportive, etc. and those who received no feedback would feel the atmosphere was unfriendly, cold, unpleasant, threatening, etc. As the results revealed, there

was a highly significant difference between the two groups with regard to their perceptions of the atmosphere, thus implying that the groups were successfully differentiated in terms of the FB and NO-FB treatment conditions previously specified.

Of particular relevance is the finding that both groups perceived a strong relationship between their performance and the atmosphere during the oral. Although the students were not asked about the nature of this relationship in terms of whether it had hindered or facilitated performance, there is reason to speculate from a consideration of the findings about attitude, learning, and evaluation that students who did not receive feedback and who perceived the atmosphere to be impersonal, etc., felt that their performance was adversely affected; while those students who received feedback and who perceived the atmosphere to be friendly, etc. felt that their performance was facilitated. However, as was mentioned previously, there were no significant differences between the marks assigned by the examiners and students to the two groups.

Since the oral examination is primarily a verbal exchange between examiners and examinees, it is natural to assume that the ability to communicate verbally would be a decisive factor in performance. While it was not surprising that the FB group rated themselves in their ability to communicate verbally no differently from the NO-FB group, it was interesting to find that both groups felt their ability to communicate verbally assisted them equally well during the oral. This is surprising because the two groups differed significantly in their perceptions of the atmosphere and anxiety level as it related to performance.

Perhaps it is appropriate at this point to speculate on an observation made during the course of the investigation. The researcher got the impression that on the whole, the NO-FB group, not receiving feedback from the examiners, tended to speak at length about the subject under consideration because, it was suspected, they felt that the answers they had given were either inadequate or incorrect. The result was that these students were observed to repeat themselves, elaborate on the topic, or ask directly whether they were correct or not, i.e., they tended to persevere. On the other hand, students in the FB group tended toward early closure, i.e., they would make comments about the topic and upon receiving feedback from the examiners, would venture little or no further information. In effect, for many of them feedback often acted as a signal that they had finished answering the question.

The attitude of many students and examiners towards examinations of any kind has typically been a negative one. Yet they recognize that examinations are indeed necessary even if only to serve as a diagnostic tool. It seems a shame that many educators fail to consider that the primary objective of education is to educate and that evaluative instruments should be subservient to this objective. The prevalent negative attitude toward examinations is a direct consequence of a failure to recognize this fact. The ongoing process that is education demands flexibility in examination procedures. It is the writer's belief that the oral examination should be used in conjunction with other types of examining techniques.

This study, has, in part, looked at some of the conditions

that influence students' perception of the oral examination. An important aspect of this perception concerns attitude. Given that examinations are necessary and given that oral examinations are desirable, what then are some of the variables that would help to produce the greatest amount of learning and at the same time not elicit negative reactions to examination?

Perhaps one answer to this dilemma is related to the amount and type of feedback. Feedback itself is necessary but not sufficient. It seems that the key to successful use of feedback lies with the manner in which it is given.

Part of the purpose of this study was to find out whether positive feedback within a supportive, helping atmosphere in comparison to a situation where feedback was withheld, would result in differential student responses toward the oral examination. The results were quite conclusive: in addition to a reported gain in knowledge and understanding, the subjects in the feedback group had attitudinal responses that were significantly more positive than those in the no-feedback group.

These findings are significant in that they provide evidence that examinations, and particularly oral examinations can be conducted in a manner that need not elicit negative attitudes while at the same time fulfilling the major objective of evaluating.

CHAPTER VII

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary and Conclusions

This study was undertaken because the writer considered that the oral examination could be a vital, evaluative instrument, which, in the past, has been avoided in favour of more expedient but not necessarily more valid ways of examining students. It is also a recognition of the importance of knowing and understanding students' perceptions of the oral situation and some of the variables that seem to influence it.

A total of 43 subjects, comprising two groups, were used in this study. Feedback was freely given to one group while in the other, feedback was withheld. Hypotheses were postulated around a number of areas concerning learning, evaluation, atmosphere, anxiety, communication, and attitudes as it related to an oral examination. Each student was given an oral examination which lasted approximately thirty minutes; it was based on an essay he had prepared. There were three examiners, the instructor in the course, in addition to two graduate students, one of whom was the researcher. After all the oral examinations were given, students' responses toward the oral were obtained through the use of a questionnaire and a semantic differential.

It was found that the Ss in the FB group tended to view the oral examination as more of a learning experience than the NO-FB subjects. Support for this conclusion was evident from an analysis of the data elicited on the semantic differential but not by the data from the questionnaire. Compared to the Ss in the NO-FB group, Ss in the FB group

also claimed that they experienced greater gain in terms of knowledge and better understanding of the implications of their essays.

Both groups viewed the oral examination as a better evaluative tool than a written examination. They reported, too, that the oral examination provided them with a greater opportunity to discuss or defend their essays and gave them a better opportunity to explain more than they would normally have had in a written examination. However, the FB group made responses that were significantly more positive than the NO-FB group.

There was no significant difference between the two groups in ratings of their level of anxiety just before the oral examination, rated ex post facto. However, Ss in the NO-FB group rated themselves as being more anxious during the oral examination than those in the FB group. Both groups felt that their anxiety adversely affected their performance, but the NO-FB more so than the FB group. An analysis of the marks assigned by examiners and students revealed no significant differences between the two groups.

Compared to Ss in the FB group, Ss in the NO-FB group perceived the atmosphere during the oral examination as being unfriendly, impersonal, threatening, etc. Furthermore, both groups felt that the atmosphere made a difference in their performance. The writer would hypothesize that the atmosphere assisted the Ss in the FB group in their performance; for those in the NO-FB group, the atmosphere probably affected their performance adversely.

The hypothesis that the FB and NO-FB group would rate themselves no differently in their ability to communicate verbally, was sustained. However, the prediction that Ss in the FB group would feel that their

ability to communicate verbally assisted them more in the oral examination than the Ss in the NO-FB group, was not upheld. Both groups rated themselves the same on this question.

On the whole, it was found that Ss in the FB group had more positive attitudinal responses toward the oral examination than those in the NO-FB group.

Inter-examiner and examiner-students reliabilities for the marks assigned on the oral examination were calculated. Correlation coefficients between the marks assigned by the examiners all exceeded .60 when the composite marks were considered. However, correlations between marks assigned by the examiners and marks assigned by the students themselves were very low, but positive. When the individual criterion grading variables were analyzed, the correlation coefficients were found to be slightly lower than that for the composite marks.

Action Recommendations

On the basis of the results of this study, it can be shown that the oral examination can be a fairly reliable instrument and if educators are to fully utilize the resources available for the evaluation of students, then the oral examination should be high on a list of priorities involving evaluative instruments. The oral examination need not be as formal as the ones used in this and other studies. The teacher may find it expedient to try a group-type oral examination in which several students are evaluated at the same time. Other teachers involved in a similar course of study can be brought in as examiners.

It is also recommended that the oral examination should constitute only a part of the total process of evaluation and can be administered

at any time during the school year. Furthermore, it is obvious that feedback, similar to that received by students in the feedback condition of this study, must be provided during an oral examination if any real benefit is to be gained.

Finally, there is reason to believe that oral examinations can be meaningfully used with students at almost all levels of the educational system. In particular, there is need to increase the oral communication skills of students at the younger age levels since so much of one's subsequent performance is conditioned by the kinds of experiences gained in earlier years.

Research Recommendations

High on a list for future research is the need to delineate and examine the nature and kinds of feedback that can be given in an oral situation and the resulting effect on learning, attitudes and some of the other variables mentioned in this study. A closely related research project would be an investigation of the actual process of communication between examiners and examinees in an oral examination. Perhaps such a study could produce some explanations for the phenomenon of early "closure" that was evident in subjects who were in the feedback condition of this study.

SELECTED REFERENCES

- Ager, N. & Weltman, J. The present structure of university examinations. Universities Quarterly, 1967, 21(3), 272-285.
- Allport, G. Attitudes. In C. Murchison (Ed.), A Handbook of Social Psychology. Worcester, Mass.: Clark University Press, 1935.
- Atiyeh, N.N. Examinations, trends and prospects. In J.A. Lauwreys, and D.G. Scanlon (Ed.), The World Year Book of Education. New York & Harcourt, Brace & World Inc., 1969.
- Barnes, E.J. & Pressey, S.L. The reliability and validity of oral examinations. School and Society, 1929, 30, 722.
- Benton, A.L., Hartman, C.H. & Sarason, I.G. Some relations between speech behavior and anxiety level. Journal of Abnormal and Social Psychology, 1955, 51, 295-297.
- Bloom, B.S. Taxonomy of educational objectives, Handbook I - Cognitive domain. New York: David McKay Co. Inc., 1956.
- Brereton, J.L. Theories of examination. In J.A. Lauwerys, & D.G. Scanlon (Ed.), The World Book of Education. New York: Harcourt, Brace & World Inc., 1969.
- Campeau, P.L. Test anxiety and feedback in programmed instruction. Journal of Educational Psychology, 1968, 59, 159-163.
- Carter, H.D. How reliable are good oral examinations? California Journal of Educational Research, 1962, XIII, 147-153.
- Chambers, W.M. Testing and its relationship to educational objectives. Journal of General Education, 1964, 16, 246-249.
- Cox, R. Examinations and higher education. Universities Quarterly, 1967, 21, 292-340.
- Cox, R. Reliability and validity of examinations. In J.A. Lauwerys & D.G. Scanlon (Ed.), The World Year Book of Education. New York: Harcourt, Brace & World Inc., 1969.
- Eisenson, J., Aver, J.J., & Irwin, G.V. The Psychology of Communication, New York: Appleton-Century-Crofts, 1963.
- Gaier, E.L. Students' perception of factors affecting test performance. Journal of Educational Research, 1962, 55, 561-566.
- Graham, G. Denmark's oral examinations. Education, 1963, 83, 306 - 309.

- Gynther, R.A. The effects of anxiety and of situational stress on communicative efficiency. Journal of Abnormal and Social Psychology, 1957, 54, 274-276.
- Hartog, P. & Rhodes, E.C. An examination of examinations. In R. Cox, Reliability and validity of examinations. The World Year Book of Education. New York: Harcourt, Brace & World Inc., 1969.
- Hassol, L., Cameron, N. & Margaret, A. Production of scattered speech via personalized distraction: An investigation of continuity in normal and schizophrenic language. American Psychologist, 1952, 7, 351-352.
- Hoyland, C.I. Reconciling conflicting results derived from experimental and survey studies of attitude change. American Psychologist, 1959, 14, 8-17.
- Insko, C.A. Theories of Attitude Change. New York: Appleton-Century-Crofts, 1967.
- Katz, D., & Stotland, E.A. A preliminary statement to a theory of attitude structure and change. In S. Koch (Ed.), Psychology: A study of science. Vol. 3, New York: McGraw-Hill, 1959.
- Kerrick, J.S. The effect of relevant and non-relevant sources of attitude change. Journal of Social Psychology, 1958, 47, 15-20.
- Kerlinger, F.N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1964.
- Kiesler, C.A., Collins, B.E., & Miller, N. Attitude Change. New York: Wiley & Sons, 1969.
- Krech, D. & Crutchfield, R.S. Theory and problems of social psychology. New York: McGraw-Hill Book Co., 1948.
- Lana, R.E. Pre-test-treatment interaction effects in attitudinal studies. Psychological Bulletin, 1959, 56, 293-300.
- Leavitt, H.J. & Mueller, R.A.H. Some effects of feedback on communication. Human Relations, 1951, IV, 410.
- Mandler, G. & Sarason, S.B. A study of anxiety and learning. Journal of Abnormal Social Psychology, 1952, 47, 166-173.
- McGuire, W.J. Nature of attitudes and attitude change. In G. Lindzey and E. Aronson (Eds.), Handbook of Social Psychology, Reading, Mass: Addison-Wesley, 1968.

- McKeachie, W.J. Procedures and techniques of teaching. In N. Sanford (Ed.), The American College, New York: 1962.
- Morrissett, I. An experiment with oral examinations. Journal of Higher Education, 1958, 29, 185-190.
- Narciso, Jr., J.C. A note on an experimental testing procedure. Journal of Educational Research, 1964, 58.
- Oppenheim, A.M., Jahoda, M. & James, R.L. Assumptions underlying the use of university examinations. Universities Quarterly, 1967, 21(3), 341-351.
- Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. The Measurement of Meaning. University of Illinois, 1957.
- Pieron, H. Examens et docimologie. In R. Cox, Reliability and validity of examinations. The World Yearbook of Education. New York: Harcourt, Brace & World Inc., 1969.
- Powell, N.J. Relation of order of appearance to mark in oral tests. Journal of Applied Psychology, 1943, 27, 104-108.
- Raymond, C.K. Anxiety and task as determiners of verbal performance. Journal of Experimental Psychology, 1953, 46, 120-124.
- Rosenberg, M. Cognitive structure and attitudinal effect. Journal of Abnormal Social Psychology, 1956, 53, 367-372.
- Shapiro, J.G. Responsivity to facial and linguistic cues. Journal of Communication, 1968, 18, 11-17.
- Snider, J.G. & Osgood, C.E. (Eds.), Semantic Differential Technique Chicago: Aldine Publishing Co., 1969.
- Solomon, D. & Rosenberg, L. Teacher-student feedback and classroom social structure. Journal of Social Psychology, 1964, 62, 197-210.
- Staubb, M. An experiment in oral testing. Modern Language Journal, 1955, 39, 232-236.
- Trimble, O.C. The oral examination - Its validity and reliability, School and Society, 1934, 39, 550-552.
- Watzlawick, P., Beavin, J.H. & Jackson, D.D. Pragmatics of Human Communication. New York: Norton & Co., 1967.

APPENDIX A

ORAL EXAMINATION QUESTIONNAIRE - PILOT STUDY

PILOT STUDY

QUESTIONNAIRE N = 37

Purpose: to determine students' reactions to, and usefulness of, the oral evaluation in an under-graduate course.

1. Do you favor oral examinations as one means of evaluation?
 - a. Yes (86%)
 - b. No (11%)
2. Did you feel that the atmosphere during the oral examination was
 - a. friendly (89%)
 - b. impersonal (11%)
3. Do you think the atmosphere made a difference in your performance on the oral?
 - a. Yes (78%)
 - b. No (22%)
4. Is this your first experience with an oral?
 - a. Yes (86%)
 - b. No (14%)

Questions 5-7 -- If you had an opportunity to take another oral, in preparation for it would you:

5. Make sure the main concepts in your essay were clear in your mind?
 - a. Yes (97%)
 - b. No (3%)
6. Make an attempt to understand the implications for teaching?
 - a. Yes (94%)
 - b. No (3%)
7. Obtain opinions about your work from other students?
 - a. Yes (56%)
 - b. No (42%)

Questions 8-9 -- Do you feel your performance on the oral was better because:

8. You could verbalize concepts better than you can write them down?

a. Yes (50%)

b. No (50%)

9. You had prepared for it?

a. Yes (44%)

b. No (56%)

Questions 10-12 -- Do you feel your performance on the oral was worse because:

10. There was little relationship between the essay and the questions you were asked?

a. Yes (11%)

b. No (83%)

11. You had not prepared for it?

a. Yes (28%)

b. No (69%)

12. You write better than you speak?

a. Yes (31%)

b. No (67%)

Questions 13-17 -- Did the oral examination help you to:

13. Re-focus your attention on your essay before the oral?

a. Yes (72%)

b. No (28%)

14. Clarify concepts that were previously vague?

a. Yes (56%)

b. No (44%)

15. See the implications of the main focus of your study?

a. Yes (81%)

b. No (19%)

16. See ways in which you would have changed your essay?
- a. Yes (78%)
 - b. No (22%)
17. Explain what you would not have normally communicated in a written examination?
- a. Yes (69%)
 - b. No (31%)

Questions 18-26 -- Do you believe that the oral will help you in the future to:

18. Prepare a better term paper?
- a. Yes (69%)
 - b. No (31%)
19. Feel less anxious on future orals?
- a. Yes (72%)
 - b. No (28%)
20. Do you feel the questions asked by the examiners were stimulating?
- a. Yes (78%)
 - b. No (19%)
21. Do you feel your experience on the oral gives you a better understanding of what goes on in an oral situation?
- a. Yes (94%)
 - b. No (6%)
22. Do you feel the questions asked by the examiners were fair?
- a. Yes (97%)
 - b. No (3%)
23. Do you believe that the mark you received on your oral was a fair assessment of what you felt you had done?
- a. Yes (69%)
 - b. No (31%)
24. Would you like an opportunity to discuss your performance on an oral?
- a. Yes (69%)
 - b. No (31%)

25. If it should be feasible in teaching, would you use an oral type of evaluation?
- a. Yes (94%)
 - b. No (6%)
26. Would you consent to having your oral videotaped for examination and research purposes? (Only the examiners to see?)
- a. Yes (61%)
 - b. No (31%)

For each question, indicate only one choice:

27. In relation to other members of the class, would you say that you were:
- a. much below average. (1%)
 - b. below average (17%)
 - c. average. (53%)
 - d. above average. (28%)
 - e. much above average. (0%)
28. Would you rather have the oral examination based on:
- a. essay alone. (42%)
 - b. essay and lectures. (8%)
 - c. essay, lectures and seminars. (47%)
 - d. lectures and seminars. (3%)
 - e. lectures alone. (0%)
29. How much do you think the oral should count for in terms of the overall evaluation of your performance in this course?
- a. 15% (39%)
 - b. 33% (42%)
 - c. 40% (14%)
 - d. 50% (6%)
 - e. 66% (0%)
30. Do you feel your anxiety during the oral:
- a. helped you on your performance. (19%)
 - b. made no appreciable difference in your performance. (25%)
 - c. affected your performance adversely. (47%)
 - d. made no difference. (8%)

31. The length of the oral should have been:

- a. 20 minutes. (36%)
- b. 30 minutes. (50%)
- c. 10 minutes. (0%)
- d. 40 minutes. (11%)
- e. longer than 40 minutes. (3%)

APPENDIX B

EXAMINER'S SELF-EVALUATION SHEETS FOR NO-
FEEDBACK AND FEEDBACK CONDITIONS

UNIVERSITY OF ALBERTA

Examiner's Name _____

Student's Name _____

EXAMINER'S SELF-EVALUATION: NO-FEEDBACK CONDITION

In Examining the S, I	Not Applicable	None of the Time	Some of the Time	Most of the Time	All of the Time
Refrained from Providing Ver- bal Feedback					
Refrained from Providing Non- verbal Feedback					
Avoided Answer- ing Questions					
Limited Inter- action to Ques- tions and Answers					
Refrained from Following up a Student's Response					

UNIVERSITY OF ALBERTA

Examiner's Name _____

Student's Name _____

EXAMINER'S SELF-EVALUATION: FEEDBACK CONDITION

In Examining the S, I	Not Applicable	None of the Time	Some of the Time	Most of the Time	All of the Time
Provided Feed- back of S					
Answered S's Questions					
Encouraged Inter- action through Non-Verbal Means					
Encouraged Inter- action through Verbal Means					
Encouraged S to Discuss Implica- tions of his Essay					
Used a Friendly an and Supportive Approach					

APPENDIX C

ORAL EXAMINATION QUESTIONNAIRE - PRESENT STUDY

NAME _____

ORAL EXAMINATION QUESTIONNAIRE

This questionnaire is part of a research study designed to assess the nature and value of oral examinations. The value of this questionnaire will in large part, depend on how frank you are in stating your opinions, feelings, and attitudes. Your individual responses will be kept strictly confidential.

We are requesting you to give your name because it will be necessary for research purposes.

Should any student wish to obtain a copy of the analysis of the data it will be readily available. The actual analysis will not be done until your final examination grades have been submitted. It will not count toward your final mark.

If you have any questions, please ask the person who has passed out the questionnaire.

DIRECTIONS: There are no "catch" questions on this questionnaire. Please read over each question and each scale very carefully. There is no time limit

Here is how you are to use the Scales:

If you feel that the concept is very closely related to one end of the scale, you should circle the number as follows:

0 1 2 3 4

or

0 1 2 3 4

If you feel that the concept is quite closely related to one or the other end of the scale (but not extremely), you should circle the number as follows:

0 1 2 3 4

or

0 1 2 3 4

If you consider the concept to be somewhere in the middle of the scale, then you should circle the scale as follows:

0	1	2	3	4
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

Sometimes you will feel as though you have had the same item before on the questionnaire. This will not be the case. Make each item a separate and independent judgement. Thank you for your cooperation.

ANSWER ALL QUESTIONS

1. To what extent do you feel you have gained in terms of knowledge as a result of having the oral examination?

0	1	2	3	4
(no gain)				(very much gain)

2. To what extent do you feel you have gained in terms of better understanding of the implications of your essay for teaching as a result of having the oral examination?

0	1	2	3	4
(no gain)				(very much gain)

3. How would you describe the atmosphere during the oral examination?

0 1 2 3 4
(impersonal) (very friendly)

4. Do you feel that the atmosphere created during the oral examination:

0	1	2	3	5
(made no differ- ence in my per- formance)				(made a tremen- dous difference in my performance)

5. Did the oral examination give you an opportunity to explain more than you would have normally communicated in a written examination?

[illegible]

6. How would you rate your anxiety level just before the oral examination?

0	1	2	3	4
(no anxiety)				(great anxiety)

7. How would you rate your anxiety level during the oral examination?

0	1	2	3	4
(no anxiety)				(great anxiety)

8. To what extent do you feel that your anxiety adversely affected your performance during the oral examination?

0	1	2	3	4
(no adverse effects)				(great adverse effects)

9. To what extent did the oral examination help you to see ways in which the essay could be improved?

0	1	2	3	4
(no help)				(great help)

10. Do you feel the oral examination was significantly different from a written examination?

0	1	2	3	4
(not significantly different)				(highly significant difference)

11. Has your motivation for learning more about your essay topic increased as a result of having the oral examination?

0	1	2	3	4
(no increase)				(tremendous increase)

12. To what extent would you view the oral examination as a learning experience?

0	1	2	3	4
(not a learning experience)				(very much of a learning experience)

13. To what extent would you view the oral examination as a better evaluative tool than a written examination?

0	1	2	3	4
(no better)				(very much better)

14. How do you rate yourself on your ability to communicate verbally?

0	1	2	3	4
(poor)				(excellent)

15. To what extent do you feel that your ability to communicate verbally, assisted you in the oral examination?

0	1	2	3	4
(no assistance)				(very much assistance)

16. To what extent do you feel the oral examination provided you with an opportunity to discuss or defend your essay?

0	1	2	3	4
(no opportunity)				(very much opportunity)

APPENDIX D

SEMANTIC DIFFERENTIAL

Name _____

SEMANTIC DIFFERENTIAL

The purpose of this study is to determine attitudes toward certain aspects of the oral examination. You are asked to rate these aspects on a series of descriptive scales. Please make your judgements on the basis of what it means to you. Make your ratings on each of the scales in order.

Here is how you are to use the scales:

If you feel that the aspect of the oral examination is very closely related to one end of the scale, you should circle number 1 or 7 as follows:

fair 1 2 3 4 5 6 7 unfair

or

fair 1 2 3 4 5 6 7 unfair

If you feel that the aspect of the oral examination is quite closely related to one or the other end of the scale (but not extremely), you should circle number 2 or 6:

severe 1 2 3 4 5 6 7 lenient

or

severe 1 2 3 4 5 6 7 lenient

If you feel that the aspect of the oral examination seems only slightly related to one side as opposed to the other side (but is not really neutral), then you should circle number 3 or 5:

strong 1 2 3 4 5 6 7 weak

or

strong 1 2 3 4 5 6 7 weak

If you consider the aspect of the oral examination to be neutral on the scale, both sides of the scale equally associated, or if the scale is completely irrelevant, unrelated to the concept, then you should circle the number 4:

hot 1 2 3 4 5 6 7 cold

*IMPORTANT

1. Be sure you check every scale. Do not omit any.
2. Never put more than one check-mark on a single scale.
3. Make each item a separate and independent judgement.

Work at a fairly high speed through this test. Do not worry or puzzle over individual items. It is your first impressions, the immediate "feelings" about the items we want. On the other hand, please do not be careless, because we want your true impressions.

Thank you for your cooperation!

SEMANTIC DIFFERENTIAL(i) ORAL EXAMINATION

1. fair	1	2	3	4	5	6	7	unfair
2. Impersonal	1	2	3	4	5	6	7	personal
3. like	1	2	3	4	5	6	7	dislike
4. hostile	1	2	3	4	5	6	7	friendly
5. rigid	1	2	3	4	5	6	7	flexible
6. effective	1	2	3	4	5	6	7	ineffective
7. democratic	1	2	3	4	5	6	7	autocratic
8. helpful	1	2	3	4	5	6	7	useless
9. depressing	1	2	3	4	5	6	7	exciting
10. good	1	2	3	4	5	6	7	bad
11. wide	1	2	3	4	5	6	7	narrow
12. active	1	2	3	4	5	6	7	passive
13. incompetent	1	2	3	4	5	6	7	competent
14. important	1	2	3	4	5	6	7	trivial

15.	attractive	1	2	3	4	5	6	7	repulsive
16.	cool	1	2	3	4	5	6	7	warm
17.	aggressive	1	2	3	4	5	6	7	unenterpris- ing
18.	successful	1	2	3	4	5	6	7	unsucces- ful
19.	insensitive	1	2	3	4	5	6	7	sensitive
20.	sad	1	2	3	4	5	6	7	happy
21.	valuable	1	2	3	4	5	6	7	worthless
22.	adaptable	1	2	3	4	5	6	7	rigid
23.	accepted	1	2	3	4	5	6	7	rejected
24.	involved	1	2	3	4	5	6	7	withdrawn
25.	insignificant	1	2	3	4	5	6	7	significant
26.	dishonest	1	2	3	4	5	6	7	honest

SEMANTIC DIFFERENTIAL

(ii) The Educational Psychology 471 Oral Examination

As a Learning Experience

1. effective	1	2	3	4	5	6	7	ineffective
2. helpful	1	2	3	4	5	6	7	useless
3. worthless	1	2	3	4	5	6	7	valuable
4. successful	1	2	3	4	5	6	7	unsuccessful
5. bad	1	2	3	4	5	6	7	good
6. significant	1	2	3	4	5	6	7	insignificant

(iii) The Educational Psychology 471 Oral Examination As A

Means of Evaluating Students

1. fair	1	2	3	4	5	6	7	unfair
2. rigid	1	2	3	4	5	6	7	flexible
3. like	1	2	3	4	5	6	7	dislike
4. insensitive	1	2	3	4	5	6	7	sensitive

5. dishonest	1	2	3	4	5	6	7	honest
6. shallow	1	2	3	4	5	6	7	deep
7. valuable	1	2	3	4	5	6	7	worthless
8. bad	1	2	3	4	5	6	7	good
9. incompetent	1	2	3	4	5	6	7	competent

(iv) Atmosphere During Educational Psychology

Oral Examination

1. pleasant	1	2	3	4	5	6	7	unpleasant
2. cool	1	2	3	4	5	6	7	warm
3. attractive	1	2	3	4	5	6	7	repulsive
4. aggressive	1	2	3	4	5	6	7	unenterpris- ing
5. passive	1	2	3	4	5	6	7	active
6. exciting	1	2	3	4	5	6	7	depressing
7. hostile	1	2	3	4	5	6	7	friendly
8. democratic	1	2	3	4	5	6	7	autocratic

9.	sad	1	2	3	4	5	6	7	happy
10.	harmonious	1	2	3	4	5	6	7	discordant
11.	silent	1	2	3	4	5	6	7	talkative
12.	supporting	1	2	3	4	5	6	7	threatening

APPENDIX E

UNROTATED FACTOR MATRIX FOR THE

CONCEPT, 'ORAL EXAMINATION'

UNROTATED FACTOR MATRIX

Scales	1		2		3	
	Communalities	Evaluation	Atmosphere	Potency		
fair-unfair	0.628	0.761	-0.114	0.144		
impersonal-personal	0.803	0.759	0.290	-0.143		
like-dislike	0.810	0.872	-0.135	-0.022		
hostile-friendly	0.725	0.632	0.362	-0.242		
rigid-flexible	0.822	0.745	0.361	0.096		
effective-ineffective	0.880	0.908	-0.224	0.071		
democratic-autocratic	0.585	0.526	0.308	-0.139		
helpful-useless	0.756	0.819	-0.258	0.094		
depressing-exciting	0.691	0.774	-0.155	-0.183		
good-bad	0.870	0.889	-0.247	-0.109		
wide-narrow	0.562	0.584	-0.025	0.164		
active-passive	0.786	0.659	0.450	0.233		
incompetent-competent	0.712	0.843	-0.006	0.020		
important-trivial	0.629	0.584	-0.537	0.002		
attractive-repulsive	0.860	0.881	-0.157	-0.221		
cool-warm	0.554	0.492	0.537	0.082		
aggressive-unenterprising	0.887	0.177	-0.062	0.881		
successful-unsuccessful	0.870	0.909	-0.205	-0.048		
insensitive-sensitive	0.434	0.531	0.082	-0.233		
sad-happy	0.771	0.864	0.116	-0.032		
valuable-worthless	0.821	0.814	-0.393	0.032		
adaptable-rigid	0.792	0.680	0.438	0.220		
accepted-rejected	0.768	0.865	0.064	0.124		
involved-withdrawn	0.681	0.782	0.151	0.131		
insignificant-significant	0.842	0.891	-0.212	0.037		
dishonest-honest	0.692	0.796	0.100	-0.135		

APPENDIX F

VARIMAX ROTATED FACTOR MATRIX FOR THE

CONCEPT, 'ORAL EXAMINATION'

VARIMAX ROTATED FACTOR MATRIX

Scales	1			2			3		
	Communalities	Evaluation	Atmosphere	Potency					
fair-unfair	0.628	0.631*	0.210	0.494					
impersonal-personal	0.803	0.384	0.789*	0.178					
like-dislike	0.810	0.742*	0.466*	0.189					
hostile-friendly	0.725	0.253	0.683*	0.276					
rigid-flexible	0.822	0.301	0.351	0.779*					
effective-ineffective	0.880	0.820*	0.316	0.303					
democratic-autocratic	0.585	0.185	0.205	0.677*					
helpful-useless	0.756	0.773*	0.178	0.338					
depressing-exciting	0.691	0.692*	0.444	0.097					
good-bad	0.870	0.829*	0.267	0.326					
wide-narrow	0.562	0.434	-0.020	0.606*					
active-passive	0.786	0.182	0.737*	0.309					
incompetent-competent	0.712	0.629*	0.401	0.390					
important-trivial	0.629	0.792*	-0.017	0.004					
attractive-repulsive	0.860	0.773*	0.451*	0.201					
cool-warm	0.554	0.005	0.627*	0.378					
aggressive-unenterprising	0.887	0.129	0.050	0.007					
successful-unsuccessful	0.870	0.814*	0.361	0.277					
insensitive-sensitive	0.434	0.358	0.541*	0.003					
sad-happy	0.771	0.568*	0.565*	0.356					
valuable-worthless	0.821	0.864*	0.138	0.225					
adaptable-rigid	0.792	0.195	0.374	0.766*					
accepted rejected	0.768	0.592*	0.427	0.458					
involved-withdrawn	0.681	0.476*	0.558*	0.312					
insignificant-insignificant	0.842	0.801*	0.296	0.326					
dishonest-honest	0.692	0.529*	0.379	0.502*					

* Denotes scores satisfying the criterion greater than 0.5

APPENDIX G

APPENDIX G

F-TEST - DIFFERENCES BETWEEN VARIANCES

F TEST - DIFFERENCES BETWEEN VARIANCES

Variable	Var1	Var2	Df1	Df2	F	P-Nondirectional
1	0.69	1.29	21.	19.	1.858	0.1705543
2	1.43	1.01	21.	19.	1.416	0.4502719
3	1.05	0.93	21.	19.	1.132	0.7911263
4	0.54	0.45	21.	19.	1.205	0.6870020
5	1.42	1.51	21.	19.	1.061	0.8905646
6	0.97	1.11	21.	19.	1.148	0.7549943
7	1.08	0.96	21.	19.	1.130	0.7937109
8	1.05	1.49	21.	19.	1.417	0.4372218
9	1.29	1.13	21.	19.	1.143	0.7735490
10	0.25	0.55	21.	19.	2.218	0.0794550
11	1.16	0.75	21.	19.	1.548	0.3426872
12	1.24	1.11	21.	19.	1.117	0.8135183
13	1.58	1.25	21.	19.	1.267	0.6080809
14	0.93	0.33	21.	19.	2.826	0.0265421
15	1.56	0.79	21.	19.	1.977	0.1405174
16	1.51	1.35	21.	19.	1.122	0.8050676

N.B. Variables 1 to 16 correspond to the questions used on the Questionnaire in Appendix C.

APPENDIX H

GRADING CRITERIA USED TO EVALUATE

STUDENTS' PERFORMANCE ON THE

ORAL EXAMINATION

ORAL EXAMINATION RATING SHEET

1. Ability to report facts clearly.

1	2	3	4	5
poor	low average	average	high average	excellent

2. Understanding of concepts in essay.

1	2	3	4	5
poor	low average	average	high average	excellent

3. Ability to present summary of essay.

1	2	3	4	5
poor	low average	average	high average	excellent

4. Ability to argue (support or refute) topics reported in essay.

1	2	3	4	5
poor	low average	average	high average	excellent

5. Ability to discuss relevance of topic for education (both theoretical and practical implications).

1	2	3	4	5
poor	low average	average	high average	excellent

APPENDIX I

t-TESTS BETWEEN MEAN SEMANTIC DIFFERENTIAL
RESPONSES TO THE CONCEPT, 'EDUC. PSY. 471 ORAL
EXAMINATION AS A LEARNING EXPERIENCE', FOR
FEEDBACK AND NO-FEEDBACK GROUPS

Scale	Combined Mean		Standard Deviation		DF	t	P-One Tail
	No-Feed-back Group	Feedback Group	No-Feed-back Group	Feedback Group			
effective-ineffective	4.09	3.70	1.88	1.73	40.	0.681	0.249
helpful-useless	4.45	3.00	1.78	1.34	40.	2.900	0.003
valuable-worthless	4.36	3.20	1.61	1.40	40.	2.427	0.009
successful-unsuccessful	4.91	3.45	1.28	1.43	40.	3.409	0.001
good-bad	4.36	3.10	1.67	1.26	40.	2.684	0.005
significant-insignificant	4.45	3.45	1.64	1.60	40.	1.957	0.028

APPENDIX J

t-TESTS BETWEEN MEAN SEMANTIC DIFFERENTIAL
RESPONSES TO THE CONCEPT, 'THE EDUC. PSY. 471 ORAL
AS A MEANS FOR EVALUATING STUDENTS' FOR FEEDBACK
AND NO-FEEDBACK GROUPS

Scale	Combined Mean		Standard Deviation		df	T	P-One Tail
	No-Feed-back Group	Feedback Group	No-Feed-back Group	Feedback Group			
fair-unfair	4.18	3.75	1.75	1.58	40.	0.817	0.209
flexible-rigid	4.36	3.20	1.43	1.63	40.	2.403	0.010
like-dislike	4.59	3.60	1.80	1.69	40.	1.792	0.040
sensitive-insensitive	3.82	3.80	1.40	1.60	40.	0.038	0.484
honest-dishonest	3.64	3.20	1.15	1.21	40.	1.170	0.124
deep-shallow	4.32	3.95	1.58	1.50	40.	0.755	0.227
valuable-worthless	4.45	3.10	1.41	1.37	40.	3.076	0.001
good-bad	4.09	2.90	1.44	1.30	40.	2.732	0.004
competent-incompetent	3.82	3.30	1.40	1.19	40.	1.255	0.108

APPENDIX K

t-TESTS BETWEEN MEAN SEMANTIC DIFFERENTIAL
RESPONSES TO THE CONCEPT, 'ATMOSPHERE DURING THE
EDUC. PSY. 471 ORAL EXAMINATION' FOR FEEDBACK
AND NO-FEEDBACK GROUPS

Scale	Combined Mean		Standard Deviation		df	T	P-One Tail
	No-Feed-back Group	Feedback Group	No-Feed-back Group	Feedback Group			
pleasant-unpleasant	3.86	2.45	1.22	1.50	40.	3.286	0.001
warm-cool	4.50	2.65	1.37	1.24	40.	4.462	0.001
attractive-repulsive	4.23	3.25	0.90	1.51	40.	2.508	0.008
aggressive-unenterprising	3.36	3.60	1.02	1.39	40.	-0.615	0.270
active-passive	3.55	2.75	1.44	1.30	40.	1.830	0.037
exciting-depressing	4.14	3.40	1.55	1.53	40.	1.512	0.069
friendly-hostile	3.50	2.65	1.12	1.28	40.	2.245	0.015
democratic-autocratic	3.86	3.45	1.29	1.40	40.	0.974	0.167
happy-sad	4.23	3.65	1.20	1.28	40.	1.472	0.074
harmonious-discordant	4.14	3.90	1.14	1.34	40.	0.603	0.274
talkative-silent	3.59	2.45	1.19	1.28	40.	2.913	0.030
supportive-threatening	4.55	3.75	1.08	1.51	40.	1.930	0.030

APPENDIX L

t-TEST BETWEEN MEAN SEMANTIC DIFFERENTIAL

RESPONSES ON THE EVALUATIVE FACTOR FOR

FEEDBACK AND NO-FEEDBACK GROUPS

Scale	Combined Mean		Standard Deviation		df	T	P-One Tail
	No-Feed-back Group	Feed-back Group	No-Feed-back Group	Feed-back Group			
fair-unfair	4.23	3.30	1.59	1.45	40.	1.917	0.031
like-dislike	4.68	3.65	1.77	1.80	40.	1.829	0.037
effective-ineffective	4.41	3.55	1.59	1.40	40.	1.811	0.038
helpful-useless	4.36	3.35	1.49	1.53	40.	2.122	0.020
exciting-depressing	4.14	3.40	1.58	1.43	40.	1.543	0.065
good-bad	4.18	3.30	1.56	1.45	40.	1.847	0.036
competent-incompetent	3.73	3.25	1.21	1.13	40.	1.282	0.103
important-trivial	3.68	3.25	1.55	1.22	40.	0.973	0.168
attractive-repulsive	4.41	3.60	1.27	1.53	40.	1.828	0.037
successful-unsuccessful	4.45	3.55	1.56	1.60	40.	1.812	0.038
happy-sad	4.23	3.50	1.44	1.53	40.	1.545	0.065
valuable-worthless	4.14	3.00	1.63	1.10	40.	2.560	0.007
accepting-rejecting	4.18	3.00	1.37	1.14	40.	2.950	0.002
involved-withdrawn	3.59	2.25	1.64	0.89	40.	3.168	0.001
significant-insignificant	4.09	3.10	1.83	1.37	40.	1.920	0.031
honest-dishonest	3.50	3.10	1.16	1.30	40	1.020	0.154

B29980